Shape Optimization Theory and Applications in Hydrodynamics

by
Onur Geçer

B.S. in Naval Architecture and Marine Engineering
Turkish Naval Academy, Istanbul-Turkey, 1999

Submitted to the Department of Ocean Engineering
and the Department of Mechanical Engineering
in Partial Fulfillment of the Requirements for the Degrees of

Master of Science in Naval Architecture and Marine Engineering

and

Master of Science in Mechanical Engineering

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2004

Author ....
Department of Ocean Engineering
August 1, 2004

Certified by..
Triantaphyllos R. Akylas, Professor of Mechanical Engineering
Department of Mechanical Engineering
Thesis Co-Advisor

Certified by..
Paul D. Sclavounos, Professor of Naval Architecture
Department of Ocean Engineering
Thesis Advisor

Accepted by..
Ain A. Sonin, Professor of Mechanical Engineering
Chairman, Departmental Committee on Graduate Studies
Department of Mechanical Engineering

Accepted by ....
Michael S.Triantafyllou
Chairman, Departmental Committee on Graduate Studies
Department of Ocean Engineering

~ This Page Intentionally Blank ~

Shape Optimization Theory and Applications in Hydrodynamics

by

Onur Geçer

Submitted to the Department of Ocean Engineering
and the Department of Mechanical Engineering
on August 13, 2004
in Partial Fulfillment of the Requirements for the Degrees of

Master of Science in Naval Architecture and Marine Engineering

and

Master of Science in Mechanical Engineering

Abstract

The Lagrange multiplier theorem and optimal control theory are applied to a continuous shape optimization problem for reducing the wave resistance of a submerged body translating at a steady forward velocity well below a free surface. In the latter approach, when the constraint formed by the boundary conditions and the Laplace's governing equation is adjoined to the objective functional to construct the Lagrangian, the dependence of the state on the control is disconnected and they are treated as independent variables; whereas in the first approach, dependences are preserved for the application of Lagrange multiplier theorem. Both methods are observed to yield identical solutions and adjoint equations. Two alternative ways are considered for determining the variation of the objective functional with respect to the state variable which is required to solve the adjoint equation defined on the body boundary. Comparison of these two ways also revealed identical solutions. Finally, a free surface boundary is included in the optimization problem and its effect on the submerged body shape optimization problem is considered. Noting that the analytical solution to the local optimization problem holds for any initial body geometry, it is therefore concluded that the above study will provide theoretical background for an efficient hydrodynamic shape optimization module to be coupled with up-to-date flow solvers currently available such as SWAN.

Thesis Advisor : Paul D. Sclavounos, Professor of Naval Architecture
Department of Ocean Engineering

Thesis Advisor : Triantaphyllos R. Akylas, Professor of Mechanical Engineering
Department of Mechanical Engineering

Acknowledgments

First of all, I would like to thank my thesis advisor Professor Paul D. Sclavounos for his continuous support and instructive guidance both on my research work and my academic studies. I would like to thank my thesis reader Professor Triantaphyllos R. Akylas from the Department of Mechanical Engineering for his additional advice and help.

Thanks to my classmate Greg Tozzi for the valuable discussions on the subject.

Funding for my study was provided by the Turkish Navy. I acknowledge their support. I am grateful to Professor Tarik Sabuncu for the inspiration and encouragement.

I am particularly indebted to my family for their love, understanding and support.

Contents

List of Figures

Chapter 1

Introduction

1.1 Background and Motivation

There is in general no analytical formula for the solution of convex optimization problems but, as with the linear problems, there are very effective methods for solving most of them. This is due to the fact that convex optimization is a generalization of linear problems. In convex optimization problems, we replace the more restrictive equality for linearity condition with an inequality which is enough to satisfy convexity.

Convex optimization problems are to be stated such that both the objective functional and the constraints are convex. If one can formulate a problem as a convex optimization problem, then it is most likely to be solved efficiently almost like the linear optimization problems. However, we can not yet surely claim that solving general convex optimization problems is a mature technology like linear programming problems. Research for various methods is still continuing actively and no consensus has emerged yet as to what the best methods are. Difficulty also arises due to the fact that recognizing convex optimization problems, or those that can be transformed to convex optimization problems can be challenging.

An optimization problem in which the objective or constraint functionals are not linear, but not known to be convex, is called a nonlinear optimization. There are no effective methods for solving the general nonlinear optimization problem. Even simple looking problems with a few variables can be extremely challenging. Methods for the general nonlinear optimization problems therefore take several different approaches, each of which involves some compromise.

In local optimization problems, the compromise is to give up seeking the optimal which minimizes the objective over all feasible points. Instead we seek a point that is only locally optimal, which minimizes the objective functional among feasible points that are near it, but is not guaranteed to have the lowest objective value. Local optimization problems can handle large scale problems and are widely applicable, since they only require differentiability of the objective and constraint functions. They are widely used in applications where there is value in finding a good point, if not the very best. In an engineering design application as we did in this study, local optimization can be used to improve the performance of a design originally obtained by manual, or other design methods.

The local optimization methods require an initial guess for the optimization variable. This starting point is critical and can greatly affect the objective value of the local solution obtained. Little information is provided about how far from globally optimal the local solution is. Therefore, local optimization methods are considered to include some level of art. Since differentiability of the objective and constraint functionals is the only requirement for most local optimization methods, formulating a practical problem as a nonlinear optimization problem is relatively straightforward. The art is in solving the problem once it is formulated. In convex optimization, the art and challenge is in problem formulation; once it is formulated as a convex optimization problem, it is relatively straightforward to solve it.

In global optimization, the exact global solution of the optimization problem is found, but the compromise is efficiency. Global optimization is used for problems with a small number of variables, where the computing time is not critical, and the value of finding the true global solution is very high.

Based on the considerations briefly discussed above, it is possible to benefit from the advantage of both approaches, if we can reduce the nonlinearity to a convex optimization problem in the application. This can be achieved in a few ways. One obvious way is to combine convex optimization with a local optimization method. Starting with a non-convex problem, we first find an approximate, but convex, formulation of the problem. By solving this approximate problem, which can be done easily and without an initial guess, we obtain the exact solution to the approximate convex problem. This point is then used as the starting point for a local optimization method, applied to the original non-convex problem.

Another approach is to compute a lower bound on the optimal value of the non-convex problem for global optimization. Two standard methods for doing this are based on convex optimization. In 'relaxation', each non-convex constraint is replaced with a looser, but convex, constraint. In Lagrangian duality approach, the convex Lagrangian dual problem is solved and a lower bound on the optimal value of the non-convex problem is obtained.

A continuous shape optimization problem in hydrodynamics will be analyzed in this study. Leaving aside the difficulty of recognizing convexity of objective and constraint functionals, Lagrange dual functional will be defined which will yield a lower bound on the optimal value of the original optimization problem. Thus, the lower bound depends on the adjoint or Lagrange multiplier value, which needs to be determined to answer the question: What is the best lower bound that can be obtained from Lagrange dual functional? Although we construct the problem as a Lagrange dual problem which is most likely to be convex independent of the original problem, uncertainty remains due to other conditions, namely constraint qualifications, to be satisfied. However, we will still obtain a very valuable inequality due to the weak duality property. This property will give the gap between the optimal value of the primal problem and the best lower bound on it. The magnitude of this duality gap will enable us to comment on the result that we will obtain based on the initial geometry that we start from.

The continuous nature of the shape optimization problem at hand and the relatively easier case governed by Laplace's equation will provide us with the flexibility to consider the problem also in terms of local optimization by simply necessitating the differentiability of the objective and constraint functionals. The result will be affected by the selected initial geometry, i.e the starting point. However, the continuous solution to the problem will be valid for any arbitrary initial geometry, which will therefore maximize the efficiency of local optimization approach to the problem. Along with the continuous Lagrange multipliers approach, the problem will be formulated also in terms of optimal control theory and the two approaches will be compared. Finally, a free surface boundary condition will be included in the problem to observe the effects on it. All these efforts are intended to form a theoretical background for an efficient hydrodynamic shape optimization routine to be coupled with the currently available state-of-the-art flow solvers such as SWAN.

1.2 Thesis Organization

The remainder of this thesis is organized as follows:

Chapter 2 gives the essential theory such as generalization of the concepts of differentials, gradients and the definition of the first variation of functionals. It states the first necessary condition for a local extremum and derives the Euler-Lagrange equations which has found extensive use for a long time for simplistic approaches to the hydrodynamic shape optimization problem.

Chapter 3 provides the theoretical background for constrained local optimization problem, leading to the Lagrange multiplier theorem.

Chapter 4 provides the theoretical background for optimal control theory, and states a simplified extension of it to be applied to our steady case.

Chapter 5 applies the optimal control theory to the shape optimization problem, derives the gradient leading to the optimal solution and also defines the adjoint equations and boundary conditions.

Chapter 6 evaluates the variation of the objective functional with respect to the state variable which is necessary for the solution of the body boundary adjoint equations derived.

Chapter 7 gives the application of the Lagrange multiplier theorem to the same problem, compares the adjoint equations and the gradient solution with the ones that are obtained by means of optimal control theory. It also briefly considers the addition of the free surface boundary condition and its effect on the submerged body shape optimization problem.

Chapter 8 concludes the thesis study.

Chapter 2

2.1 Introduction to Variational Theory

The principles of orthogonality and projection theorem, expressed in various ways, form the basis of the optimization principle. In spite of the large variety of norm definitions available for minimum or least-norm problems, many optimization problems can not be formulated directly in these terms and therefore, optimization of more general objective functionals needs to be considered. Yet, geometric interpretation and the theory obtained from minimum norm problems provide insight to the more general optimization problems [1,7], also considered in this thesis.

Before aiming at the general optimization problems, we first need to generalize the concepts of differentials, gradients etc. to normed spaces. By using these tools, it is possible to relate the variational theory of general optimization to the familiar theory in the finite dimensional case.

2.2 The First Variation in the Calculus of Variations

In the calculus of variations, integral functionals of the form

$$J(x) = \int_{t_1}^{t_2} f[t, x(t), \dot{x}(t)]dt$$

are considered on the interval $[t_1, t_2]$, where x is a member of some functional space. One then seeks the extremals x of the functional J, such that $J(x) - J(\tilde{x})$ has the same sign for all $\tilde{x}$ in a neighborhood around x.

The requirements of the particular problem at hand determine a neighborhood. For example, a *strong extremum* is given when we consider x as an element of the space $D[t_1, t_2]$ of continuous functions on $[t_1, t_2]$ with the norm

$$\|x\| = \sup_{t \in [t_1, t_2]} |x(t)|$$

Here, the term *norm* defines an abstraction of our usual concept of length.

Definition. If a real-valued function defined on a vector space X maps each element x in X into a real number $\|x\|$, the vector space is called a *normed linear vector space* and the real number $\|x\|$ is called the norm of x.

In the above definition, **sup** stands for the *least upper bound* or *supremum* of a set of real numbers, here $[t_1, t_2]$, bounded above by the smallest real number y such that $x \le y$ for all $x \in [t_1, t_2]$.

A *weak extremum* arises when we choose x from a space $D^1[t_1, t_2]$ of continuously differentiable functions with the norm defined as

$$\|x\| = \sup_{t \in [t_1, t_2]} |x(t)| + \sup_{t \in [t_1, t_2]} |x'(t)| .$$

In either case, a neighborhood of x is given by all those functions $\tilde{x}$ such that $\|x - \tilde{x}\| < \varepsilon$ for some $\varepsilon > 0$. Since there are fewer functions in a weak neighborhood of x, it is easier for x to be an extremal in the second case.

Let us now define the variation of the functional J.

Definition. The variation of δJ of the functional J is the linear part, if exists, of the increment:

$$\Delta J[h] = J[x + h] - J[x]$$

that is, $\delta J[h]$ is the linear functional, which differs from $\Delta J[h]$ by an infinitesimal of order higher than 1 relative to $\|h\|$.

Variation, like the differential of a real valued function, is defined at a specific point x of the domain of the functional and it is a functional on the tangent space at x.

Before laying the foundation for a theory of extreme values of functionals and deriving the first necessary condition for a relative minimum of a functional, let us generalize the concept of the derivative to functionals that are defined on normed linear spaces over $\mathbb{R}$, or at least on open subsets thereof.

2.3 Gateaux and Frechet Differentials

Let T be a transformation defined on an open domain D in a normed space X and having range in a normed space Y. Here, transformation is an extension of the familiar notion of ordinary functions. It is simply a function defined on one vector space X, taking values in another vector space Y. And a special case of this situation is that in which the space Y is taken to be the real line. In this case, a transformation from a vector space X into the space of real (or complex) scalars is said to be a *functional* on X.

Returning to our definition, we call δT (x;h) the **Frechet differential** or **Strong differential** of T at x with increment h if there is a $\delta > 0$ such that for all $h \in X$, $\|h\| < \delta$,

$$T(x+h) - T(x) = \delta T(x;h) + \varepsilon_1(h)$$

where δT (x;h) is a linear functional of h and where $\lim_{\|h\| \to 0} [\varepsilon_1(h) / \|h\|] = 0$. Note that T(x;h) is called a linear functional of h if it is additive, that is, if T(x;h+k) = T(x;h) + T(x;k) for all h,k $\in X$.

A second, somewhat weaker concept of the differential of a functional is the Gateaux or weak differential. Again, let X be a vector space, Y a normed space, and T a transformation defined on a domain $D \subset X$ and having range $R \subset Y$. δT (x;h) is called the **Gateaux** or **Weak differential** of T at x with increment h if there is a $\delta > 0$ such that for all $h \in X$, $\|h\| < \delta$,

$$T(x + \alpha h) - T(x) = \delta T(x;h) + \varepsilon_2(h)$$

where δT (x;h) is a linear functional of h and where $\lim_{\alpha \to 0}[\varepsilon_2(\alpha h)/\alpha] = 0$. (α real).
The concept of the Gateaux differential is somewhat weaker than the concept of the Frechet differential, since in the case of Frechet differential, $\varepsilon_1(h)$ has to tend to zero uniformly in h, while in the case of Gateaux differential, $\varepsilon_2(h)$ only has to tend to zero along each $h \in X$. It is called weak differential because its definition requires no norm on X; hence, properties of the Gateaux differential are not easily related to continuity. When X is normed, a more satisfactory definition is given by the Frechet differential.

We can see that for fixed $x \in D$ and h regarded as variable, the Gateaux differential defines a transformation from X to Y. In the particular case where T is linear, we have, by linearity explained above, $\delta T(x;h) = T(h)$.

In the case where Y is the real line, the transformation reduces to a real-valued functional on X. Thus, if *f* is a functional on X, the Gateaux differential of *f*, if it exists, is

$$\delta f(x;h) = \frac{d}{d\alpha} f(x+\alpha h) \Big|_{\alpha=0}$$

and for each fixed $x \in X$, $\delta f(x;h)$ is a functional with respect to the variable $h \in X$.

The Gateaux differential generalizes the concept of directional derivative that we are familiar with in finite dimensional space. The following example will make similarity more obvious.

Example. Let X=E^n and let $f(x) = f(x_1, x_2, ..., x_n)$ be a functional on E^n having continuous partial derivatives with respect to each of its variables. Then the Gateaux differential of f is

$$\delta f(x;h) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} h_i$$

This is a more general abstract expression of the well known directional derivative;

$$D_{\vec{h}} f = \vec{\nabla} f . \vec{h}$$

of the function f, in the direction of h.

Before closing our discussion of more general differentials, let us give one more example of the frequently used Gateaux differentials, use of which will be made in the following parts.

Example. Let $X = C[t_1, t_2]$ and let $f(x) = \int_{t1}^{t2} g[x(t), \dot{x}(t), t] dt$ where it is assumed that g_x and $g_{\dot{x}}$ exists and continuous with respect to x', x and t. Then,

$$\delta f(x;h) = \frac{d}{d\alpha} \int_{t1}^{t2} g[x(t) + \alpha h(t), \dot{x}(t) + \alpha \dot{h}(t), t] dt \; \Big|_{\alpha=0}$$

by replacing the order of integration with differentiation and expanding in power series, differential takes the form:

$$\delta f(x;h) = \int_{t1}^{t2} g_x(x, \dot{x}, t) h(t) dt + \int_{t1}^{t2} g_{\dot{x}}(x, \dot{x}, t) \dot{h}(t) dt$$

2.4 First Necessary Condition for a Local Extremum

We are now ready to extend the well known technique of minimizing a function of a single variable to a similar technique based on more general differentials briefly defined in the previous section. Therefore, we can obtain the analogs of the first necessary condition for local extrema.

Let f be a real valued functional defined on a subset Ω of a normed space X . Ω, on which the extremum problem is defined denotes the space of competing functions, in other words, it is the admissible set. We assume that Ω admits a linear space of admissible variations κ.

Definition. For a given space of competing functions $\Omega \in X, \kappa \in X$ is called a space of admissible variations of Ω if, for all $x \in \Omega, h \in \kappa, (x+h) \in \Omega$.

If f(x) assumes a relative minimum at $x_0 \in \Omega$ relative to elements $x \in \Omega$, then it is necessary that

$$f(x_0 + h) - f(x_0) \geq 0$$

for all $h \in \kappa$, where κ is a space of admissible variations of Ω, so long as $\|h\| < \delta$ for some $\delta > 0$.

If $f(x)$ possess a Gateaux variation $\delta f(x;h)$ at x_0, Then,

$$f(x_0 + h) - f(x_0) = \delta f(x;h) + \varepsilon(h)$$

for all $h \in X$ for which $\|h\| < \delta$ for some $\delta > 0$.

From the two equations given above, we have

$$\delta f(x;h) + \varepsilon(h) \geq 0$$

for all $h \in \kappa$, for which $\|h\| < \delta$. We choose an arbitrary $h_0 \in \kappa$ for which $\|h_0\| < \delta$. Then, because κ is a linear space, we have $\alpha h_0 \in \kappa$ for all $\alpha \in \Re$, and if $\|\alpha\| \leq 1$, we have $\|\alpha h_0\| < \delta$. By considering the homogeneity of $\delta f(x;h)$,

$$\alpha \delta f(x_0; h_0) + \varepsilon(\alpha h_0) \geq 0. \qquad (\text{ for all } |\alpha| \leq 1.)$$

If $0 < \alpha \leq 1$,

$$\delta f(x_0; h_0) + \frac{\varepsilon(\alpha h_0)}{\alpha} \geq 0.$$

and if $-1 \leq \alpha < 0$,

$$\delta f(x_0; h_0) + \frac{\varepsilon(\alpha h_0)}{\alpha} \leq 0.$$

From the definition of the Gateaux differential $\lim_{\alpha \to 0}[\varepsilon_2(\alpha h)/\alpha] = 0$. We obtain $\delta f(x_0; h_0) = 0$.

Since h_0 was an arbitrary element of κ with $\|h_0\| < \delta$, we obtain immediately the **necessary condition ;**

$$\delta f(x_0; h) = 0$$

for all $h \in \kappa$, for which $\|h\| < \delta$ for some $\delta > 0$.

Theorem. If the functional $f(x)$, which is presumed to possess a Gateaux variation at $x_0 \in \Omega \subset X$, assumes a relative minimum (maximum) in Ω at $x = x_0$ and if Ω admits a linear space of admissible variations κ, then it is necessary that

$$\delta f(x_0; h) = 0 \qquad \text{for all } h \in \kappa$$

$f(x)$ does not need to be defined on the entire space Ω as long as it is defined in an open subset $Y \subset X$ that contains x_0.

A point at which $\delta f(x_0; h) = 0$ for all h is called a *stationary point;* hence, the above theorem states that extrema occurs at stationary points. A similar result holds for a local extremum of a functional defined on an open subset of a normed space, the proof being identical for both cases.

The simplicity of the above theorem is of great utility. Much of the calculus of variations can be regarded as a simple application of this one result. Many interesting problems are solved by careful identification of an appropriate vector space X and some algebraic manipulations to obtain the differential.

2.5 The Euler-Lagrange Equation

Let us find the function x on the interval [t1,t2] which minimizes an integral functional of the form

$$J = \int_{t1}^{t2} f[x(t), \dot{x}(t), t]dt$$

We assume that the function f is continuous in $x, \dot{x}$ and t, and it has continuous partial derivatives with respect to x and $\dot{x}$. We seek a solution in the space D[t1,t2]. For the simpler case, we assume that the end points x(t1) and x(t2) are fixed. This will further restrict the admissible set, the class of functions within which we seek the extremum.

Starting with a given admissible vector x, we consider vectors of the form x+h that are admissible, too. The class of vectors $h \in \kappa$ is called admissible variations. In our problem, the class of admissible variations becomes the subspace of D[t1,t2], with the elements (functions) that vanish at *t1* and *t2*. The necessary condition for the extremum problem is derived in the previous section and stated again as below,

$$\delta f(x_0; h) = 0 \qquad \text{for all } h \in \kappa$$

The differential of J is,

$$\delta J(x;h) = \frac{d}{d\alpha} \int_{t1}^{t2} f[x + \alpha h, \dot{x} + \alpha \dot{h}, t]dt \ \Big|_{\alpha=0}$$

or as we have presented previously as in one of the most commonly used format;

$$\delta J(x;h) = \int_{t1}^{t2} f_x(x, \dot{x}, t)h(t)dt + \int_{t1}^{t2} f_{\dot{x}}(x, \dot{x}, t)\dot{h}(t)dt$$

for all $h \in \kappa$ as a point of departure. We denote;

$$\int_{t1}^{t} f_x[s, x(s), \dot{x}(s)]ds = \phi(t)$$

applying integration by parts to the first term of the differential above, we obtain

$$\int_{t1}^{t2} f_x[t, x(t), \dot{x}(t)]h(t)dt = h(t)\phi(t) \Big|_{t1}^{t2} - \int_{t1}^{t2} \phi(t)\dot{h}(t)dt$$

Since h(t1) = h(t2) = 0 and since ϕ is continuous, the first term on the right hand side of the above equation vanishes and we have consequently;

$$\int_{t1}^{t2} \dot{h}(t)\{f_{\dot{x}}[t, x(t), \dot{x}(t)] - \phi(t)\} dt = 0$$

for all $h \in \kappa$. Lemma of Dubois-Reymond will enable us to transform this condition still further.

Lemma of Dubois-Reymond. If $\alpha(t)$ is continuous in [t1,t2] and $\int_{t1}^{t2} \alpha(t)\dot{h}(t)dt = 0$ for every $h \in D[t1, t2]$ with h(t1) = h(t2) = 0, then $\alpha(t) \equiv c$ in [t1,t2] where c is a constant.

Proof. Let c be the unique constant satisfying $\int_{t1}^{t2} [\alpha(t) - c]dt = 0$ and let

$$h(t) = \int_{t1}^{t} [\alpha(s) - c]ds$$

Then,

$$\int_{t1}^{t2} [\alpha(t) - c]\dot{h}(t) = \int_{t1}^{t2} \alpha(t)\dot{h}(t)dt - c[h(t2) - h(t1)] = 0$$

and hence $\alpha(t) \equiv c$.

By making use of the Lemma of Dubois-Reymond we can say that

$$f_{\dot{x}}[t, x(t), \dot{x}(t)] - \phi(t) = c$$

And replacing the term $\phi(t)$ with its definition we can state first necessary condition for a relative extremum:

Theorem. For x to yield a relative extremum for the integral $J = \int_{t1}^{t2} f[x(t), \dot{x}(t), t]dt$, it is necessary that there be a constant c such that the integro-differential equation

$$f_{\dot{x}}[t, x(t), \dot{x}(t)] = \int_{t1}^{t} f_x[s, x(s), \dot{x}(s)]ds + c$$

is satisfied by $x = x_0(t)$ for all $t \in [t1, t2]$ except for the points where $\dot{x}_0$ has a jump discontinuity. The above equation is called the Euler-Lagrange equation in integrated form.

By noting that $\int_{t1}^{t} f_x[s, x(s), \dot{x}(s)]ds = \phi(t)$ again, we can write the Euler-Lagrange equation in its differential form;

$$f_x[t, x(t), \dot{x}(t)] - \frac{d}{dt}\{f_{\dot{x}}[t, x(t), \dot{x}(t)]\} = 0$$

Euler Lagrange equation and isoperimetric problems, in which one is required to find an extremum of an integral with a subsidiary condition (such as the volume or the waterline area is constant) has found many applications in the theoretical study of the problem of determining the ship form of minimum wave resistance. However, these studies could achieve only results for simple geometries or approximations (such as thin ship theory) and they do not correspond to real ships. Therefore, as we direct the interested reader to a good collection of all these efforts [5,9], it should be noted that the above stated theory of generalized differentials will be utilized to form the background of the more general optimization routines which, later in this study, will be put into practical applications.

Chapter 3

3.1 Theoretical Background for Constrained Optimization

Due to the difficulty which is met to define a practical problem involving convexity of the functionals, the local theory of Lagrange multipliers provides a wider applicability and a general convenience for most of the optimization problems. Although the theory itself becomes simpler and more elegant with convex functionals involved in global constrained optimization problem, the above consideration makes the local theory better known and it is most commonly applied in practical sense.

The principles of both the local and global theories are essentially the same. Optimization problems with inequality constraints are almost identical for both of the theories. Particularly, some important difference is observed for equality constraints[7]. As one can see in the following sections of this chapter, the difference for equality constraints arise from the fact that we base the local theory on acceptable approximations in the primal space X, the space where we define our objective functional (or we define it on a subset of the primal space X; mostly defined as D). An auxiliary operation is then required to relate these approximations to the constraint space Z, the space on which the constraint is mapped. These two spaces are related to each other because the objective function optimization in the space X is achieved within the constraint limits defined in the space Z.

In fact, as we will see later, we need to relate the approximations in the primal space X with the space Z*, the space where we will define Lagrange multipliers. A linear operator is then used in the development of the local theory, namely adjoint operator, which enables us to transfer optimization results in X* back to Z*. Here X* represents the space of all bounded linear functionals on X and is called the *normed dual* or *dual* of X.

Usually an effective algorithm can be built on the dual approach, which depends in an essential way on convexity and therefore on global theory. The theory of conjugate convex functions [2,3,1] furnishes a fairly complete answer to the question of how to construct the duality. Although the main results of conjugate function theory apply only to optimization problems of convex type, being attracted by the convenience that they provide, we can see many applications also to non-convex problems [13]. Some of these concern the derivation of necessary conditions for optimality. Others arise because, in the course of a proper algorithm, a non-convex problem is approximated locally at each step by a convex problem, as a way of defining the optimal direction for gradient based optimization problems.

We need to mention a little more about the dual spaces here, because as will be seen in the following pages, interrelations between a space and its corresponding dual, that is the space consisting of all bounded linear functionals on the original space, plays a very important role in the optimization theory defined on normed spaces. Dual space provides a 'dual' setting for the optimization problem defined in the original (primal) space, as commonly denoted as X. It creates the alternative in a sense that if the primal problem is a minimization problem, the dual problem becomes a maximization problem. Lagrange dual functional gives a lower bound on the optimal value of the optimization problem, which depends on the values of Lagrange multipliers. In that case, we try to reach the best lower bound that can be obtained from the Lagrange dual functional.

Lagrange dual problem is a convex optimization problem, since the objective to be maximized is concave (it is point wise infimum of a family of affine functions depending on the Lagrange multiplier values) and the constraint is convex. This is the case whether or not the primal problem is convex.

Optimal values of both of these objective functions are equal and the solution of one of the problems leads to the solution of the other. Dual spaces are also essential for the development of the concept of a gradient and they also provide the setting for Lagrange multipliers, which become the fundamental tool for the constrained optimization problem. Dual spaces also play a role that is analogous to the inner product defined in Hilbert space.

Therefore we can develop results that are the extension of the projection theorem solution of the minimum norm problems into arbitrary normed linear spaces [7].

One brief explanation about the notation might be useful for the remaining part of this text. As a general rule, functional on a normed space X, linear or not, are denoted by f, g, h etc., that is analogous to the common function denotation. However, for the specific steps of development, certain bounded linear functionals will be shown as x_1^*, x_2^*, etc. as elements of the dual space X*. The value of the linear functional $x^* \in X^*$ at the point $x \in X$ is denoted by x*(x) or by the more symmetric notation <x, x*>.

General Lagrange multiplier theorem forms the basis for all of the constrained optimization problems, and it was devised by Lusternik. This theorem will be given in the next section. Generalized inverse function theorem, a difficult analysis underlying this theorem, is contained, but not given in detail at the beginning of this part. For the practical-aimed purpose of this study, it is considered to be enough to understand the statement of the theorem. However for a complete proof, one is directed to the relevant reference [8].

3.2 Definition of a Regular Point

We will base our derivation of generalized Lagrange multiplier theorem onto the inverse function theorem. Before introducing this theorem, the notion of the regular point will be discussed first.

Definition. Define a continuously Frechet differentiable transformation; T, from an open set D in a Banach space X into a Banach space Y. If $x_0 \in D$ is such that $T'(x_0)$ maps X onto Y; the point x_0 is said to be a ***regular point*** of the transformation T.

The existence of a regular point, i.e. being able to map a certain point defined on a space onto another space by both a transformation and the Frechet differential of that transformation provides the flexibility to define a neighborhood region that satisfies the solution as will be seen in the Generalized Inverse function theorem definition. This is the acceptable approximation in the primal space X, as we have mentioned in the beginning of this chapter, which the local theory of equality-constrained optimization is based on. We will later relate this approximation to the Z* space by means of adjoint operators. Regular point definition also enables the use of gradients, and therefore instead of the variations defined on actual surfaces, we can deal with the variations in the tangent space, and look for the availability of a stationary case, which gives an extremum.

One can also think of the definition of regular point in terms of the ordinary regular function definitions. However, the statement of this similarity should not cause any misinterpretation other than providing a better understanding. As we know, if a function can be expanded in power series near $x = x_0$, as given below for example in Taylor series;

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(x_0)}{n!}(x - x_0)$$

then, it is said to be a regular function at $x = x_0$. In order the above expansion to be defined, all derivatives of $f(x)$ must exist at $x = x_0$. The general regular point definition, however, requires the existence of first derivative, (Frechet differentiability of the transformation) which is an extension of the classical optimization theory that is based on the first derivative.

The optimization problem is defined in Banach Space. Determination of the most convenient space for the problem is beyond the scope of this study. However, we must briefly note the advantage of using this space by giving the definition first:

Definition. If every Cauchy sequence from a normed linear vector space X has a limit in X, X is called a 'complete normed linear vector space', or ***Banach space.***

Cauchy sequence can be explained within the concept of convergence. In a normed linear space, an infinite sequence of vectors $\{x_n\}$ is said to converge to a vector x, if the sequence $\{\|x - x_n\|\}$ of real numbers converges to zero. In this case, we write $x_n \to x$.

Based on this, a sequence $\{x_n\}$ in a normed space is said to be a Cauchy sequence if $\|x_n - x_m\| \to 0 \;\; as \;\; n, m \to \infty$ .i.e., given $\varepsilon > 0$, there is an integer N such that $\|x_n - x_m\| < \varepsilon$ for all n, m> N.

Normed spaces in which every Cauchy sequence is convergent (complete) are of particular interest in analysis; in complete spaces, it is possible to identify convergent sequences without explicitly identifying their limits. Compared to the other problems defined in incomplete spaces, the principal advantage of defining the optimization problem in Banach space is that in such a problem, we seek an optimal vector that is maximizing/ minimizing a given objective. In this case, we often construct a sequence of vectors, each member of which is superior to the preceding members, that is closer to the optimum result. The desired optimal vector is then the limit of the sequence. In order that scheme be effective, there must be a convergence test which can be applied when the limit is unknown. If the underlying space is complete, i.e. if it is a Banach space, Cauchy criterion for convergence meets this requirement.

3.3 Generalized Inverse Function Theorem

The proof of the theorem will not be given here but a clear description of the statement will be given briefly. That will enable us to derive the generalized Lagrange multiplier theorem in the next section.

Theorem. If x_0 is a regular point of a transformation T mapping the Banach space X into the Banach space Y, then there is a neighborhood $N(y_0)$ of the point $y_0 = T(x_0)$ (such as a sphere centered at y_0) and a constant K such that the equation T(x) = y has a solution for every $y \in N(y_0)$ and the solution satisfies $\|x - x_0\| \le K \|y - y_0\|$.

The existence of a regular point, i.e. being able to map a certain point defined on a space onto another space by both a transformation and the Frechet differential of that transformation provides the flexibility (acceptable approximation) to define a neighborhood region that satisfies the solution.

3.4 Necessary conditions for the Local Theory of Optimization

Let us know give the necessary conditions for an extremum of an objective f that is subject to the constraint $H(x) = \theta$ (i.e. null vector), where f is a real valued functional on a Banach space X (primal space) and H is a mapping from X into a Banach space Z (constraint space).

Lemma. If f and H are continuously Frechet differentiable in an open set containing the regular point x_0, and if f is assumed to achieve a local extremum that is subject to $H(x) = \theta$ at the point x_0; then $f'(x_0)h = 0$ for all h satisfying $H'(x_0)h = \theta$.

Assume specifically that the local extremum is a local minimum so that we can define the sign of the variation from the extremum explicitly. A transformation is later defined with the real-valued functional that satisfies the constraint mapping such as $T(x) = (f(x), H(x))$. This transformation maps the x values on primal space X onto a space defined as the vector product of R and Z; i.e. $T : X \to R \times Z$. This is because the functional f takes values on the real-values axis R and the constraint mapping H(x) maps the elements of the primal space onto the constraint space Z. So, the elements of the new transformation are defined on such a vector product space that they take real values that also satisfy constraint mapping on the Z plane. We introduce an increment of variation h such that $H'(x_0)h = \theta,\ f'(x_0)h \neq 0$. In other words we equate the Frechet differential of the transformation H to null vector for the minimum value of x_0; $\delta H(x_0;h) = H'(x_0)h$. Here, $H'(x_0)$ is the Frechet derivative of H and it is a bounded linear operator from primal space X to the constraint space Z.

We assume that for all the increments of variation h, which equates the Frechet differential of the constraint equation to null vector, the Frechet differential of the real-valued functional F is not equal to zero, i.e. $\delta f(x_0;h) = f'(x_0)h \neq 0$. Since f is a real valued functional, Frechet derivative of this functional can be called the gradient of f at x_0. It is a linear operator from primal space X to its dual X*.

$T'(x_0) = (f'(x_0), H'(x_0))$ will also be onto the product space, $T': X \to R \times Z$. This is because of the definition above that assumes x_0 to be a regular point of the constraint mapping H. The linear operator $H'(x_0)$ maps the point x_0 onto Z as does $H(x_0)$ by regularity and $f'(x_0)h$ is not equal to zero but takes some real value by the initial assumption. It would then follow, by the inverse function theorem that for any $\varepsilon > 0$, there exists a vector x and $\delta > 0$ with $\|x - x_0\| < \varepsilon$ such that $T(x) = [\ f(x_0) - \delta\ ,\ \theta\ ]$. This contradicts the initial assumption that x_0 is a local minimum.

Let us visualize the above result geometrically in the Primal space X as given by Luenberger [7].

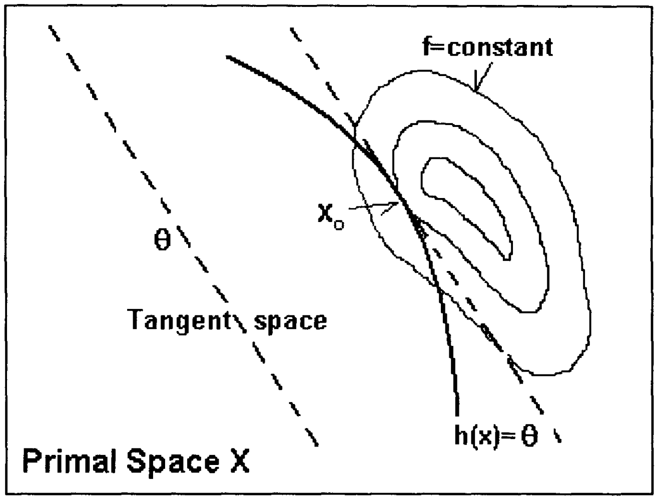

Figure 3-1: Geometric visualization of the necessary condition in Primal space.

In the figure above, contours of constant f and the constraint surface for a single functional constraint h(x)=0 are drawn in the Primal space. The above result is expressed here with the introduction of the tangent space of the constraint surface. It is a subspace of the primal space X and it comprises the set of all vectors h (admissible variations) for which $H'(x_0)h = \theta$. As we have given the definition previously, it is also called the nullspace of $H'(x_0)$. When we translate it to the optimal point x_0, it is regarded as the tangent of the surface $N = \{x : H(x) = \theta\}$ near x_0. The geometric equivalent of the necessary condition given in the above lemma is that f is stationary at the optimal point x_0 with respect to the admissible variations in the tangent plane.

3.5 Local Theory of Optimization with Equality Constraints

We have showed that an extremum f is stationary with the regularity provided, (Frechet differential of f is equal to zero) with respect to the variations in the tangent plane. In other words, we have replaced the constraint with its linearized form and the extremum f revealed to be stationary with respect to variations in this linearized constraint. Now, we can proceed to represent the Lagrange multiplier theorem by introducing the duality relations between the range and null space of a linear operator and its adjoint.

Theorem. If the continuously Frechet differentiable functional f has a local extremum that satisfies the constraint $H(x) = \theta$ at a regular point x_0, then, there exists an element $z_0{*} \in Z{*}$ such that the Lagrangian functional

$$L(x) = f(x) + z_0 {*} H(x)$$

is stationary at the point x_0, i.e.;

$$f'(x_0) + z_0 {*} H'(x_0) = \theta.$$

Proof. The set of all vectors h, for which the constraint equation, $H'(x_0)h = \theta$, is satisfied is called the *nullspace* of $H'(x_0)$. With a geometric description of the statements above, it is the tangent space at x_0.

From the Lemma in part 2.4, we know that $f'(x_0)h = 0$ for all h which satisfies $H'(x_0)h = \theta$. This necessitates that $f'(x_0)$, the linear operator from primal space X to its dual X*, is orthogonal to h, i.e. to the nullspace of the linear operator $H'(x_0)$.

Definition: Closed Set

The following steps that we will follow through the optimization process will require the range of the linear operator to be closed. Therefore, we have better explain the definitions of these concepts.

A point $x \in X$ is said to be a *closure point* of a set P, which is a subset of X, if, given a $\varepsilon > 0$, there is a point $p \in P$ satisfying $\|x - p\| < \varepsilon$. The collection of all closure points of P is called the *closure* of P and is denoted by $\overline{P}$. In other words, a point x is a closure point of P, if every sphere centered at x contains a point of P. It is obvious that P becomes a subset of $\overline{P}$. A set P is said to be *closed* if $P = \overline{P}$.

One example to the closed set definition that is easy to visualize in mind is the unit sphere, consisting of all points x with $\|x\| \leq 1$. A single point is a closed set, the empty set and the whole set X are also closed (as well as open).

Definition: The Range of a Transformation

If a transformation T maps the space X into Y; $T: X \rightarrow Y$, the collection of all vectors $y \in Y$ for which there is an $x \in X$ (or a subset of it) with $y = T(x)$ is called the *range* of T.

Since we have approximated the constraint equation with its linearized version, we will be focused on the linear transformations (linear operators) in our derivation of Lagrange Multiplier theorem. Therefore, we can define the range of the linear operator $H'(x_0)$: $X \rightarrow Z$ as follows: It is the collection of all vectors in Z for which there is an $x \in X$ (That will be our solution, i.e., the vector that gives the optimum solution; x_0) with $H'(x_0) = \theta$.

As we have stated previously, for all h satisfying $H'(x_0)h = \theta$, f is stationary at x_0. Therefore, the space of all h vectors, i.e., nullspace of $H'(x_0)$, represents the range of this linear operator. h is the tangent space or linearized version of the constraint space (the space composed of all vectors h such that $H'(x_0)h = \theta$; nullspace). It can be represented as a linear surface in three dimensions, or simply as a line in two dimensions. This linear property enables the range to be closed. This is because, the collection of all the closure points still remain on the tangent space and the closure of the tangent space equals the tangent space. Therefore, instead of creating variations on the constraint space, we can define variations on the tangent space, and still fulfill the closed property of the range.

Closed range is required for the optimization process because the variations are intended to remain in the constraint space (or as we did here, it remains in its linearized form).

Now that we know the range of $H'(x_0)$ is closed, let us proceed to define the duality relations between the range and nullspace of an operator and its adjoint. Before introducing this relation, let us briefly summarize the use of adjoint operators.

3.6 Adjoint Operators and Their use in optimization problems

As we have introduced above for our problem of generalizing the optimization problem, constraints in many optimization problems can be replaced by its linearized version. A linear operator, or an adjoint operator is then used in order to enable us to transfer optimization results in X* back to Z*. Here X* represents the space of all bounded linear functional on X and is called the *dual* of X. Z* is the space of all bounded linear functionals on the constraint space Z, it is the dual of Z. Due to their convenient mechanism for describing the orthogonality and duality relations, adjoint operators are of importance in the definition of the optimization problem.

Definition. Let X and Z be normed linear spaces and let $A \in B(X,Z)$. Here B(X,Z) denotes the normed space of all bounded linear operators from the normed space X into the normed space Z. The adjoint operator $A^*: Z^* \to X^*$ is defined by the following equation:

$$A^* z^*(x) = z^*(Ax)$$

Given a fixed $z^* \in Z^*$, the quantity $z^*(Ax)$ is a scalar for each $x \in X$ and is therefore, it is a functional on X. Furthermore, by the linearity of z* and A, it follows that this functional is linear. With the following equation:

$$|z^*(Ax)| \leq \|z^*\| \|Ax\| \leq \|z^*\| \|A\| \|x\|$$

We can see that this functional is bounded and is thus an element x* of X*. Then we define A*z*=x*. The adjoint operator A* is unique and linear. The relation between an operator and its adjoint is illustrated below:

$$X \rightarrow (A) \rightarrow Z$$
$$Z^* \rightarrow (A^*) \rightarrow X^*$$

After this brief review of the definition of adjoint operators, let us now give the relations between range and nullspace, in which adjoint operators play an important role due to their convenience for describing the orthogonality and duality relations.

Theorem 1. Let X and Z be normed spaces and let $A \in B(X, Z)$. Then the orthogonal complement of the range of a linear operator is equal to the nullspace of its adjoint as denoted below;

$$[\Re(A)]^{\perp} = N(A^*)$$

Proof. Let us remember the definitions and the spaces on which the range and nullspace of linear operators are defined again.

The collection of all vectors $z \in Z$ for which there is an $x \in X$ with z=Ax is the range of the linear operator A. Orthogonal complements of range of A consists of elements $z^* \in Z^*$ that are orthogonal to every vector $z \in Z$ in the range of A. In our definition above, the set $\{z^* : A^* z^* = \theta\}$ corresponding to the linear operator A* is the nullspace of A*. And it is a subspace of Z*.

Let $z^* \in N(A^*)$ and $z \in \Re(A)$. Then z=Ax for some $x \in X$. The following equation;

$$z^*(z) = z^*(Ax) = A^* z^*(x) = 0$$

shows that $N(A^*)$ is a subset of $[\Re(A)]^{\perp}$. If we assume that $z^* \in [\Re(A)]^{\perp}$, then for every $x \in X$, $z^*(Ax) = 0$. This is equivalent to $A^* z^*(x) = 0$ and therefore, $[\Re(A)]^{\perp}$ is a subset of $N(A^*)$. This is possible only if $[\Re(A)]^{\perp} = N(A^*)$.

The dual equivalent of the above theorem is also valid, given the range of the linear operator A is closed. i.e.;

$$\Re(A^*) = [N(A)]^{\perp}$$

In the equivalent dual equation, since the adjoint of the linear operator A is defined by: $A^* : Z^* \to X^*$, the range of A* now becomes the collection of the vectors $x^* \in X^*$ for which there is a $z^* \in Z^*$ with A*z*=x*. The nullspace of the linear operator A is simply composed of the elements of the set $\{x : Ax = \theta\}$, and it is a subspace of the primal space X.

The orthogonal component of the null space, therefore, consists of elements $x^* \in X^*$ that are orthogonal to every vector $x \in X$ in the nullspace of the linear operator A.

If we let $x^* \in \Re(A^*)$, then A*z*=x* for some $z^* \in Z^*$. And for any x that is an element of $N(A)$; the following equation would be valid;

$$x^*(x) = 0 = A^* z^*(x) = z^*(Ax)$$

Then, x* is an element of the orthogonal component of the $N(A)$ and it follows that $\Re(A^*)$ is a subset of $N(A)^{\perp}$. In order to prove that $N(A)^{\perp}$ is also a subset of $\Re(A^*)$, which leads to the result of their equality, we need to invoke extension form of Hahn-Banach theorem. This complicated proof is not provided in this study. However, we must understand the general idea behind this theorem that leads to the convenient duality relations between the range and nullspace of an operator and its adjoint.

Extension form of the Hahn-Banach theorem extends the projection theorem to problems having non-quadratic objectives. This enables the simple geometric interpretation of the projection theorem to be preserved for more complex problems. Projection theorem, simplified in three-dimensional space, states that the shortest line from a point to a plane is the perpendicular from the point to the plane. This simple, well-known result finds direct applications in spaces of higher dimensions, as is the case with Hilbert space. The extension form serves as an appropriate generalization of the projection theorem from Hilbert space to normed spaces, such as Banach Space [1,7].

3.7 Lagrange Multiplier Theorem

Before applying the duality relations we stated in part 3.6, let us now go back and repeat the theorem we will invoke for the constrained optimization problem.

Theorem. If the continuously Frechet differentiable functional f has a local extremum that satisfies the constraint $H(x)=\theta$ at a regular point x_0, then, there exists an element $z_0{}^* \in Z^*$ such that the Lagrangian functional

$$L(x)=f(x)+z_0{}^*H(x)$$

is stationary at the point x_0, i.e.;

$$f'(x_0)+z_0{}^*H'(x_0)=\theta.$$

We have also showed in section 3.5 that $f'(x_0)$, the linear operator from primal space X to its dual X*, is orthogonal to h, i.e. to the nullspace of the linear operator $H'(x_0)$.

According to the duality relations between the range and nullspace of an operator and its adjoint that we have explained in part 3.6;

If we let $x^* \in \Re(A^*)$, then A*z*=x* for some $z^* \in Z^*$.

That is, in our terms for the problem; Since the range of $H'(x_0)$, (i.e. the linear operator A) is closed,

$$f'(x_0) \in \Re\left[H'(x_0)^*\right]$$

then, there is some $z_0^* \in Z^*$ so that A*z*=x*, i.e.:

$$H'(x_0)^* z_0^* = f'(x_0)$$

that which proves the stationary situation of our Lagrangian functional at x_0. Expressed in the common notation:

$$f'(x_0) + z_0^* H'(x_0) = \theta$$

The second term in this expression is the composition of linear transformations. Based on the regularity definition, just like the constraint equation $H(x_0)$, the linearized version of the constraint equation $H'(x_0)$ also maps the point x_0 from primal space X onto the constraint space Z. The composition of adjoint of this linear operator, $H'(x_0)^*$, with the Lagrange multiplier $z_0^* \in Z^*$ maps the optimization problem onto the dual of the primal space, X*. This is the normed Banach space where we have established the extension form of the projection theorem for minimum norm problems. From the application point of view, minimum norm problems are formulated in a dual space in order to guarantee the existence of a solution. This approach has provided to us a convenience to relate the null space and range of a linear operator and its adjoint.

In almost most of the optimization problems, Lagrange multipliers are often treated in a naïve way as a convenient set of constants multiplied with the constraint equations. However, as described above, we should not consider them as individual Lagrange multipliers. In fact, they are an entire Lagrange multiplier vector defined in an appropriate dual space. The misunderstanding or simplification of the vectoral concept is usually caused by the need to represent the problem in finite dimensions. In fact, if we have a single functional equation as the constraint, on geometric grounds, the Lagrange multiplier turns out to be a scalar, i.e. a constant that multiplies the constraint. However, this should not cause a misinterpretation of the more general concept defined in vector spaces.

Lagrange multipliers are commonly used in the constraint optimization problem, because of the convenient mechanism introduced for the duality relations between the range and nullspace of an operator and its adjoint. Reviewing our arguments above, we first showed that at an extremum,f is stationary with respect to variations in the tangent space, in other words, we have replaced a nonlinear constraint with its linearized version and defined the variation on it. The Lagrange multiplier then followed from the adjoint relations.

Chapter 4

4.1 Theoretical Background for Optimal Control Theory

With the general theorem given on Lagrange Multipliers, this part contains the derivation of the optimal control theory, which, to some extent can be considered as an extension or an application of the Lagrange multiplier theory. The structure of the optimal control theory will be paid some attention and that will enable us to generalize the similar optimization methods found in the literature with different names. After the introduction of the optimal control theory, all these presumably different methods will be observed to follow the same idea. This will be more mature once the theory is followed by some applications. After all, optimal control theory will provide an additional insight into the Lagrange multipliers as a more general application to the optimization problem.

4.2 Basic Necessary Conditions

Considering an interval $[t_0, t_1]$ of the real line, we introduce a dynamic system of the form

$$\dot{x}(t) = f\left[x(t), u(t)\right]$$

that is described by a set of differential equations. Here x(t) is an n-dimensional 'state vector', u(t) is an m-dimensional 'control vector', and f is a mapping of $E^n \times E^m$ into E^n. The above dynamic system produces a vector-valued function x, when supplied with an initial state $x(t_0)$ and a control input function u.

In addition to the dynamic equation and the initial condition, we are given an objective functional to be minimized of the form

$$J = \int_{t_0}^{t_1} l(x, u)\,dt$$

and also a finite number of terminal constraints

$$g_i\left[x(t_1)\right] = c_i, \quad i = 1, 2, \ldots\ldots, r$$

or given in the vector form as

$$G\left[x(t_1)\right] = c$$

The functions l and g are assumed to possess continuous partial derivatives with respect to their arguments. The control problem is then that of finding the pair of functions (x,u) minimizing the objective functional J while satisfying the dynamic system and the terminal constraint.

In order for the general variational theory to be applied, the control problem is formulated in various spaces each with its own particular advantage. The most natural approach is to consider the problem as one formulated in $X \times U$ and to treat the dynamic condition and the terminal constraint as constraints connecting u and x. We can therefore apply the general Lagrange multiplier theorem to these constraints with separate Lagrange multipliers defined for these constraint equations individually. The optimization process then seeks to replace the Lagrange multipliers in terms of only one multiplier and by means of the initial and terminal conditions, the pair of functions (x,u) are found.

Another approach is to note the fact that once the control vector u is specified, the dynamic condition uniquely defines the state vector x with the set of differential equations and therefore, we only need to select the control vector u. The problem can then be regarded as formulated in the space of admissible control functions, U. Lagrange multiplier theorem needs only to be applied to the terminal constraints. Another approach is to view the problem in the space of admissible trajectories, X, which consists of the functions x resulting from the application of a given control u. We consider them as continuous n-dimensional functions on the interval $\left[t_0, t_1\right]$. The problem can also be defined in the finite dimensional space E^r which corresponds to the vector of terminal constraint. As we have previously stated each of these approaches has theoretical advantages for the purpose of deriving necessary conditions and practical advantages for the purpose of developing computational procedures for obtaining solutions. Our approach will be in the $X \times U$ space.

The differential equation of the dynamic system with the initial condition $x(t_0)$ is equivalent to the integral equation

$$x(t) - x(t_0) - \int_{t_0}^{t} f\left[x(\tau), u(\tau)\right] d\tau = \theta$$

that can be abstractly stated as a transformation A;

$$A(x,u) = \theta$$

A is a mapping from $X \times U$ into X. The Frechet differential of A is continuous under our initial assumption and it is given by the formula

$$\delta A(x,u;h,v) = h(t) - \int_{t_0}^{t} f_x h(\tau) d\tau - \int_{t_0}^{t} f_u v(\tau) d\tau$$

for $(h,v) \in X \times U$.

The terminal constraint is a mapping from X into E^r, and its Frechet differential is given by

$$\delta G(x;h) = G_x h(t_1)$$

The transformations A and G define the constraints of the problem, and their regularity can be proved. We can now give the basic necessary conditions satisfied by the solution to the optimal control problem.

Theorem. Let x_0, u_0 minimize

$$J = \int_{t_0}^{t_1} l(x,u) dt$$

subject to the initial, terminal and dynamic conditions

$$x(t_0) = fixed,\ G\left[x(t_1)\right] = c,\ \dot{x} = f(x,u)$$

and assume that the regularity conditions are satisfied. Then there is an n-dimensional vector valued function $\lambda(t)$ and an r-dimensional vector μ such that for all $t \in \left[t_0, t_1\right]$

(4.1) $$-\dot{\lambda}(t) = \left[f_x{}'\{x_0(t), u_0(t)\}\right]\lambda(t) + l_x{}'\{x_0(t), u_0(t)\}$$

(4.2) $$\lambda'(t) f_u\{x_0(t), u_0(t)\} + l_u\{x_0(t), u_0(t)\} = \theta$$

(4.3) $$\lambda(t_1) = G_x'\{x_0(t_1)\}\mu$$

The proof of the basic necessary conditions follows from the Lagrange Multiplier theorem given in the third Chapter, which is briefly given below.

If the continuously Frechet differentiable functional J has a local extremum under the constraints composed of the dynamic system and the terminal condition at the regular pair of functions (x_0, u_0), Lagrange Multiplier Theorem yields immediately the existence of the elements λ *and* μ such that the Lagrangian functional is stationary with respect to the variation of both the state and the control variables independently, i.e.

$$\int_{t_0}^{t} l_x(x_0, u_0)h(t)dt + \int_{t_0}^{t} d\lambda'(t)\left\{h(t) - \int_{t_0}^{t} f_x(x_0, u_0)h(\tau)d\tau\right\} + \mu' G_x\{x_0(t_1)\}h(t_1) = 0$$

$$\int_{t_0}^{t} l_u(x_0, u_0)v(t)dt - \int_{t_0}^{t} d\lambda'(t)\int_{t_0}^{t} f_u(x_0, u_0)v(\tau)d\tau = 0$$

for all $(h, v) \in X \times U$. Based on the terminal condition, we may take $\lambda(t_1) = \theta$. Integrating the first equation above by parts, we have

$$\int_{t_0}^{t_1} l_x(x_0, u_0)h(t)dt + \int_{t_0}^{t_1} d\lambda'(t)h(t) + \int_{t_0}^{t_1} \lambda'(t) f_x(x_0, u_0)h(t)dt + \mu' G_x h(t_1) = 0$$

This equation holds for all continuous h, therefore it particularly holds for all continuously differentiable h vanishing at t_0. Integrating the second term by parts, we have for such functions

$$\int_{t_0}^{t_1}\left\{l_x(x_0, u_0)h(t) - \lambda'(t)\dot{h}(t) + \lambda'(t) f_x(x_0, u_0)h(t)\right\}dt = 0$$

It follows that λ is differentiable on the interval (t_0, t_1) and the equation 4.1 holds. Integrating the variation of the above Lagrangian equation with respect to U by parts, equation 4.2 holds. Now by changing the boundary condition on $\lambda(t_1)$ from $\lambda(t_1) = \theta$ to $\lambda(t_1) = G_x{}'\mu$ to account for the jump that is necessary for the continuity, equation 4.3 holds to preserve the continuity of λ. In this general formulization of the optimal control theory, the necessary conditions 4.1, 4.2 and 4.3 along with the original dynamic, initial and terminal conditions form a complete system of equations; 2n first order differential equations, 2n boundary conditions, r terminal constraints and m instantaneous equations from which $x_0(t),\ \ \lambda(t),\ \ \mu,\ \ and\ u_0(t)$ can be found.

4.3 A Simpler Extension of the Optimal Control Theory

Based on the Lagrange Multiplier Theorem with its wide range of applicability, let us have a look at a simpler variation of the optimal control theory which leads to the reduced gradient of a cost functional. This brief theoretical introduction will be later applied to a practical application in the following section.

In most of the optimization applications, we prefer defining the problem in a Banach space that is partitioned, or in other words, composed of the vector product of state space and control space denoted as $X = X \times U$. As we have previously mentioned, some other definition of spaces can also be considered each with its own advantage of formulization and computational reasons. In our case, X is the state space and U is the control space. The variables in the control space U are also called design variables or design parameters. As we have introduced in Chapter 3, according to the theorem of Lagrange Multipliers, the mapping H introduces an equality constraint $H(x,u) = \theta$ and it is a mapping from the primal space X into the constraint space Z. It can be a set of algebraic equations, ordinary differential equations or partial differential equations, here represented in its simpler vector form. In most of the applications, the constraint commonly takes the form of a boundary value problem.

In terms of the control theory notation, H is the constraint mapping which is composed of the constraints A and G. As we have previously shown, A contains the dynamic system differential equation and the initial state, whereas G is the constraint imposed by the terminal state.

With this most general notation introduced, we know, from the theory of Lagrange multipliers that, if the continuously Frechet differentiable objective functional J has a local extremum under the constraint H(x,u)=0 at the regular values of the pair of functions (x,u), then, there exists an element $\lambda^* \in Z^*$ such that the Lagrangian functional is stationary, i.e.

$$J_x(x_0) + \lambda_0 * H_x(x_0) = 0 .$$

Since we have combined the constraint in just one vector form of a mapping H, we would not need to define another extra adjoint multiplier; μ, which is needed in the more general optimal control theory formulation to deal with the terminal constraint. The simplicity of the problem arises from the time-independence of the state and control variables, i.e. we would not need extra steps to integrate by parts the time variation of the Lagrange multiplier in this example. Within this context, equations 4.1 and 4.2 appear directly as the application of the Lagrange Multiplier theory, and one can be substituted into another to provide an explicit term to replace the Lagrange multiplier.

As in the case of the transformation A, the constraint mapping H(x,u)=0 uniquely determines the state for given controls. It connects the state variable and control variable with the set of constraint equations. When the constraint is adjoined to the objective functional to construct the Lagrangian functional, the dependence of the states on the controls are disconnected and they are treated as independent variables.

The time-independent form of the equations 4.1 and 4.2 follow from the above intuition and the stationary condition of the Lagrangian in this context which finds itself many practical applications one of which will be formulized in the following chapter.

(4.4) $$J_x(x_0,u_0)+\lambda_0 * H_x(x_0,u_0)=0$$

(4.5) $$J_u(x_0,u_0)+\lambda_0 * H_u(x_0,u_0)=0$$

When the constraint H(x, u) can be viewed as the governing equation from which the states x can be uniquely determined for a given set of control variables u, the above equations can be substituted one into another to solve for the Lagrange multiplier under the conditions that the constraint H is regular at the optimal point x_0 and the Jacobian H_x is invertible.

$$\lambda_0 * = -J_x(x_0,u_0)\left[H_x(x_0,u_0)\right]^{-1}$$

Substituting the Lagrange multiplier into the equation 3.5, we have

$$J_u(x_0,u_0)-J_x(x_0,u_0)\left[H_x(x_0,u_0)\right]^{-1} H_u(x_0,u_0)=0$$

The above necessary condition is stated at the local extremum point. On the way to the optimum solution, the iterative approach provides the intermediate values of the Lagrange multiplier and the gradient of the Lagrangian approaches to stationary with the constraint satisfied. In other words;

$$J_x(x,u)+\lambda * H_x(x,u)=0$$

and this yields;

$$\lambda = -J_x(x,u)\left[H_x(x,u)\right]^{-1}$$

The gradient of the Lagrangian with respect to the control variable is given by;

$$\frac{\partial L}{\partial U} = J_u(x,u)-J_x(x,u)\left[H_x(x,u)\right]^{-1} H_u(x,u)$$

The above equation is the gradient of the cost functional with respect to the control variable at constant values of the constraint, H. It is also called the reduced gradient in the literature. In other words, given as the boundary conditions of the problem, we keep the constraint constant. i.e.,

$$\delta H = H_x \delta x + H_u \delta u = 0$$

which gives us;

$$\delta x = -H_x^{-1} H_u \delta u$$

Thus, the variation of the objective functional becomes;

$$\delta J = J_u \delta u + J_x \delta x$$

$$= (J_u - J_x H_x^{-1} H_u)\delta U$$

or

$$\frac{\partial J}{\partial u} = J_u - J_x H_x^{-1} H_u$$

This result is identical to the previous equation that we have obtained. Therefore, the combination of the first two necessary conditions of the optimal control theory given previously holds for this simplistic approach to the problem, and in fact the optimal solution is obtained with the reduced gradient of the objective functional with respect to the control variable subject to the constraint.

Chapter 5

5.1 Application of Optimal Control Theory to the Shape Optimization Problem

The following hydrodynamic shape optimization problem is first introduced by Ragab[12]. Although, in his study, the problem is considered to be an adjoint approach as an alternative to the gradient-based numerical optimization techniques, it lacks a detailed reasoning of the steps followed. It is therefore intended in this part to formulate the problem within the principles of the optimal control theory.

The sketch below gives the flow domain where the optimization problem is defined.

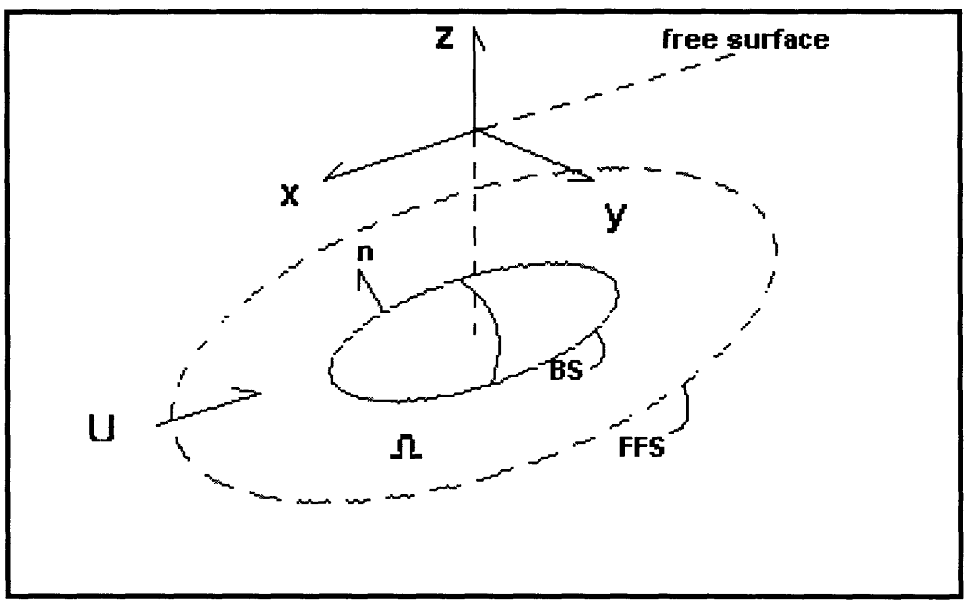

Figure 5-1: Sketch of the flow domain for a deep submergence shape optimization.

A three dimensional body is translating with a steady forward velocity U well below a free surface. We locate the body at a certain depth so that the free surface boundary condition is excluded in this simple example. The problem is defined relative to a translating coordinate system. Since the body is translating with a steady forward speed, a steady flow is defined in the opposite direction of the translation far upstream of the body with a velocity of U. The flow domain Ω is bounded by the body surface (BS) and the far-field surface (FFS).

We would like to optimize the shape of the body for a certain objective function. Therefore, our control (u) is the perturbation of the geometry, denoted with θ (a set of geometric parameters to define the body surface which is to be optimized) and for the convenience of defining the state (x) of the flow, we choose the velocity potential ϕ as our state variable.

The constraint H(u,x) appears as a set of boundary conditions given as;

(B.C.1) $$C(u,x) = \nabla^2 \phi = 0$$ (in Ω)

(B.C.2) $$B(u,x) = \frac{\partial \phi}{\partial n} + \vec{U}.\vec{n} = 0$$ (on BS)

(B.C.3) $$A(u,x) = \phi = 0$$ (on FFS)

The set of constraints given above forms a constraint space that consists of the boundary conditions. Let us consider a general objective function in the form of

$$J(u,x) = \int_{BS} f \, ds$$

The objective function is a function of both the state and the control variable. In order to form the Lagrangian by adjoining the constraint to the objective function, we can either choose to form a combination of all of the constraint equations or in an alternative way, we can treat them as individual constraints. Following the latter, each individual constraint equation that is defined on a certain boundary will then be transformed into the dual of the primal space by means of its own adjoint multiplier. It will later be possible to express the individual adjoint multipliers on certain boundaries of the fluid domain in terms of only one adjoint multiplier. This is because we need only one adjoint multiplier for a problem that is constrained with a zero-order equation, i.e. no dynamic condition, and no need for a terminal constraint to be coupled with its own adjoint multiplier.

In the formulation of the general control theory given in the previous chapter, we have defined the constraint space as one that is composed of the transformations A and G, where A is a combination of the initial state and the dynamic condition and G is the constraint imposed by the terminal constraint. That is why, we have defined two elements: λ *and* μ, where the second term, μ, that is the adjoint multiplier of the terminal constraint, has been related to the first one by the necessary condition 4.3 to provide continuity.

By adjoining the constraint equations each with its adjoint multiplier, let us define the Lagrangian as

$$L(u,x) = \int_{BS} f\,ds + \int_{\Omega} \lambda C\,d\Omega + \int_{BS} \beta B\,ds + \int_{FFS} \alpha A\,ds$$

Following the reasoning that leads to the necessary conditions 4.4 and 4.5, we implement the stationary condition of the Lagrangian by disconnecting the dependence of the state and control and treating them as independent variables. Before this step, we should distribute the constraint on the whole fluid domain imposed by the continuity equation to the boundaries of the fluid, one of which, the body boundary, will later be considered as the space that we are defining our optimization problem. For this aim, we should apply Green's theorem to the continuity condition valid throughout the whole fluid domain as below

$$\int_{\Omega} \lambda \nabla^2 \phi\, d\Omega - \int_{\Omega} \phi \nabla^2 \lambda\, d\Omega = \int_{\substack{boundaries \\ of\ the\ fluid\ domain}} \left(\lambda \frac{\partial \phi}{\partial n} - \phi \frac{\partial \lambda}{\partial n} \right) ds$$

Leaving the first term on the left hand side of the equation alone,

$$\int_{\Omega} \lambda C d\Omega = \int_{\Omega} \lambda \nabla^2 \phi d\Omega = \int_{\Omega} \phi \nabla^2 \lambda d\Omega + \int_{\substack{boundaries \\ of\ the\ fluid\ domain}} \left(\lambda \frac{\partial \phi}{\partial n} - \phi \frac{\partial \lambda}{\partial n} \right) ds$$

The boundaries of the fluid domain consists of the body surface (BS) and the far-field surface (FFS) and the last term on the right hand side of the above equation can be separated into these parts and then can be combined with the same surface integrals in the Lagrangian equation as given below;

$$L(u,x) = \int_{BS} f ds + \int_{\Omega} \phi \nabla^2 \lambda d\Omega + \int_{BS} (\beta B + \lambda \frac{\partial \phi}{\partial n} - \phi \frac{\partial \lambda}{\partial n}) ds + \int_{FFS} (\alpha A + \lambda \frac{\partial \phi}{\partial n} - \phi \frac{\partial \lambda}{\partial n}) ds$$

Once we have constructed the Lagrangian functional, the dependence of the states on the controls are disconnected and they are treated as independent variables. Following the necessary condition 4.4, we differentiate the Lagrangian with respect to the state and obtain the adjoint multipliers on each boundary.

$$\frac{\partial L(u,x)}{\partial \phi} = 0$$

$$= \int_{\Omega} \delta\phi \nabla^2 \lambda d\Omega + \int_{BS} (\frac{\partial f}{\partial \phi} \delta\phi + \beta \frac{\partial B}{\partial \phi} \delta\phi + \lambda \frac{\partial \delta\phi}{\partial n} - \delta\phi \frac{\partial \lambda}{\partial n}) ds + \int_{FFS} (\alpha \frac{\partial A}{\partial \phi} \delta\phi + \lambda \frac{\partial \delta\phi}{\partial n} - \delta\phi \frac{\partial \lambda}{\partial n}) ds$$

The above equation gives us the adjoint multipliers β *and* α in terms of λ, which relates state variable with the control variable. At the same time, adjoint equations are obtained for each separate part of the fluid domain as given below.

<u>ADJOINT EQUATIONS:</u>

(1) $$\nabla^2 \lambda = 0 \qquad \text{(in } \Omega\text{)}$$

(2) $$\frac{\partial \lambda}{\partial n} = \frac{\partial f}{\partial \phi}$$ (on BS)

(3) $$\lambda = 0$$ (on FFS)

With the above equation (2) included in the BS integration, the following steps lead to the definition of the adjoint β in terms of λ:

$$\beta \frac{\partial B}{\partial \phi} \delta\phi + \lambda \frac{\partial \delta\phi}{\partial n} = 0$$

where $B = \frac{\partial \phi}{\partial n} + \vec{U}.\vec{n}$. Therefore changing the order of differentiation;

$$\beta \frac{\partial}{\partial n}(\frac{\partial \phi}{\partial \phi})\delta\phi + \lambda \frac{\partial \delta\phi}{\partial n} = 0$$

this gives us the relation between β and λ;

(4) $$\beta = -\lambda$$

Similarly, considering the Far-field surface integration with the adjoint equation (3) included;

$$\alpha \frac{\partial A}{\partial \phi} \delta\phi - \delta\phi \frac{\partial \lambda}{\partial n} = 0$$

Where $A = \phi$. Therefore, we can easily obtain the relation between α *and* λ as;

(5) $$\alpha = \frac{\partial \lambda}{\partial n}$$

Comparing the general optimal control theory given in Chapter-4 with our application in the steady case here, we can say that since there is no time-dependence, the adjoint equations fulfill the smoothness condition. Otherwise, the terminal constraint would need to be adjoined with another multiplier, μ, and the relation between the dynamic condition coupled with the initial condition and the terminal condition were to be determined with the necessary condition 4.3 such that at the terminal point, the continuity of the Lagrange multiplier λ can still be satisfied.

In the steady case that we consider in this application, dynamic condition simply equals to zero due to the time invariance of the state variable ϕ. Therefore, we would no longer need to include a first order differential equation along with its initial condition. Due to the time invariance property of the problem, we are not given a terminal constraint to be satisfied, either. However, in spite of the simplicity of the appearance, the constraint equation still has a very important role of relating the state and control variables. Continuity of the adjoint multiplier comes by itself without necessitating a condition between the terminal constraint and the dynamic condition due the above stated property of the problem which makes the application more practical.

After determining the Lagrange multiplier functions, we can make us of the condition that the Lagrangian function is stationary with respect to the variation of the control variable as given in Equation 4.5. The adjoint equations that we have derived are valid also for the variation with respect to the control variable; θ. Therefore, we can plug in the adjoint equations for each part of the fluid domain that the problem is defined.

$$L(u,x) = \int_{BS} f\,ds + \overbrace{\int_{\Omega} \phi \nabla^2 \lambda \, d\Omega}^{=0,(Eq.1)} + \int_{BS} (\overbrace{\beta}^{=-\lambda,(Eq.4)} B + \lambda \frac{\partial \phi}{\partial n} - \phi \frac{\partial \lambda}{\partial n})\, ds + \int_{FFS} (\overbrace{\alpha}^{=\frac{\partial \phi}{\partial n},(Eq.5)} \underbrace{A}_{=\phi} + \overbrace{\lambda}^{=0,(Eq.3)} \frac{\partial \phi}{\partial n} - \phi \frac{\partial \lambda}{\partial n})\, ds$$

The above equation then simplifies to the following;

$$L(u,x) = \int_{BS} (f + \lambda \frac{\partial \phi}{\partial n} - \phi \frac{\partial \lambda}{\partial n} - \lambda B)\, ds$$

Let us now apply Green's theorem for the familiar terms on the right hand side of the above equation to further simplify it by making use of the boundary conditions and the adjoint equations before taking the variation with respect to the control variable.

$$L(u,x) = \int_{BS} (f - \lambda B)\, ds + \int_{BS} (\lambda \frac{\partial \phi}{\partial n} - \phi \frac{\partial \lambda}{\partial n})\, ds$$

$$= \int_{BS} (f - \lambda B)\, ds + \int_{\Omega} (\lambda \overbrace{\nabla^2 \phi}^{=0,(B.C.1)} - \phi \overbrace{\nabla^2 \lambda}^{=0,(Eq.1)})d\Omega - \int_{FFS} (\overbrace{\lambda}^{=0,(Eq.3)} \frac{\partial \phi}{\partial n} - \overbrace{\phi}^{=0,(B.C.3)} \frac{\partial \lambda}{\partial n})\, ds$$

Therefore, what remains in the Lagrangian equation is an integral over the body boundary;

$$L(u,x) = \int_{BS} (f - \lambda B) ds$$

Now that we have adjoined the adjoints into our Lagrangian equation, we can calculate the variation of this equation with respect to the control as given by the Equation 4.5 which will become stationary once the local optimum is reached.

$$\frac{\partial L(u,x)}{\partial \theta} = \frac{\partial}{\partial \theta} \int_{BS} (f - \lambda B) ds$$

In the above equation, the term λ needs to be calculated by means of the adjoint Equation (2). In the following chapter, we will cover the formulation to determine the variation of the objective functional with respect to the flow variable where the objective functional is an explicit function of the flow variable. Let us now calculate the above expression giving the variation of the Lagrangian with respect to the control (geometry) variable. We will denote the geometry variable with a function $\alpha(u)$ defined on the body boundary. It is a function of the geometric parameters (control variables), which we consider to be the curvilinear coordinates to define the surface shape for simplicity and we will follow the same notation in the following chapter for the ease of formulization.

The variation is composed of two terms. The first one is caused by the change of the incremental surface element and the second term is by the variation of the terms in the integration sign due to the variation of the geometry while we keep the state (flow) variable constant.

$$\frac{\partial}{\partial\theta}\int_{BS}(f-\lambda B)\,ds=\int_{BS}(f-\lambda B)\frac{\partial}{\partial\theta}(ds)+\int_{BS}\left(\frac{\partial f}{\partial\theta}-\lambda\frac{\partial B}{\partial\theta}\right)ds \tag{5.1}$$

For calculating the term $\frac{\partial}{\partial\theta}(ds)$, let us consider an infinitesimal change of geometry on the body surface in the direction of the normal vector.

$$-\varepsilon\,\alpha\,\vec{n}$$

We define the incremental piece ds as an arc

$$ds=R\,d\theta$$

where R is the radius of curvature and $d\theta$ is the incremental angle seeing the small piece of the body surface.

As we have defined previously, α is an arbitrary geometry function defined on the boundary of the surface. The change in the incremental element due to the above defined change of the geometry in the normal direction is

$$R\,d\theta-\varepsilon\,\alpha\,d\theta=(R-\varepsilon\,\alpha)\,d\theta=R\,(1-\varepsilon\,\frac{\alpha}{R})\,d\theta$$

By substituting ds in the above equation we can calculate the variation as follows;

$$\frac{1}{\varepsilon}(ds^{\varepsilon}-ds)=\frac{1}{\varepsilon}[\ (1-\varepsilon\,\frac{\alpha}{R})\,ds-ds\ ]$$

$$\frac{1}{\varepsilon}(ds^{\varepsilon}-ds)=-\frac{\alpha}{R}$$

With the above variation of the incremental surface element, the first term of the equation (5.1) gives the following:

$$\int_{BS}(f-\lambda B)\frac{\partial}{\partial\theta}(ds)=\int_{BS}\frac{\alpha}{R}(f-\lambda B)\,ds$$

However, the body boundary condition $B(u,x)=\dfrac{\partial \phi}{\partial n}+\vec{U}.\vec{n}=0$ is equal to zero as given in the problem statement. Therefore, the first term of Equation (4.1) simplifies to

(5.2) $$\int_{BS}(f-\lambda B)\frac{\partial}{\partial \theta}(ds)=\int_{BS}\frac{\alpha}{R}f\,ds$$

The second term in the same equation will be calculated next. Since we have disconnected the state and control variables in the formulation of the problem, no perturbation of the state (flow) variable will be considered but we will perturb only the geometry to calculate the variations of the terms given below once again.

$$\int_{BS}\left(\frac{\partial f}{\partial \theta}-\lambda\frac{\partial B}{\partial \theta}\right)ds$$

The second term in the above equation demands the variation of the no penetration boundary condition. Let us start with this term. We will use the total velocity potential for the ease of the following derivation which will enable us to include a steady forward speed motion buried in this term.

$$B(u,x)=\frac{\partial \phi}{\partial n}+\vec{U}.\vec{n}=\frac{\partial \Phi}{\partial n}=0$$

The variation of the above expression due to the variation of the geometry is given by;

(5.3) $$\frac{\partial B}{\partial \theta}=\frac{1}{\varepsilon}\left[\,\nabla\Phi(u^{\varepsilon}).n^{\varepsilon}-\nabla\Phi(u).n\,\right]$$

In order to represent the incremental change in the normal vector of the body surface, a simple diagram is given below.

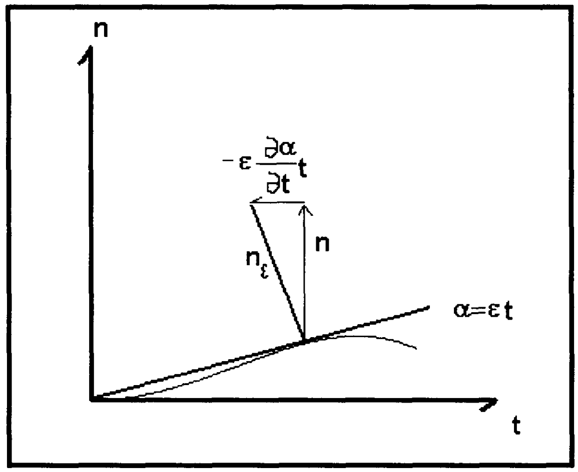

Figure 5-2: Representation of the change of the normal vector to the body surface.

Here, n and t represents the normal and the tangential coordinates respectively before the perturbation is applied. For the case of a three-dimensional object, we have two tangential coordinates and a normal coordinate to define the body surface. Based on the above diagram, n^{ε} is given with the following;

$$n^{\varepsilon} = \vec{n} - \varepsilon \frac{\partial \alpha}{\partial \xi}\vec{\xi} - \varepsilon \frac{\partial \alpha}{\partial \eta}\vec{\eta} + O(\varepsilon^2) = \vec{n} - \varepsilon \sum_{t=\xi,\eta} \frac{\partial \alpha}{\partial t}\vec{t} \tag{5.4}$$

$\Phi(u^{\varepsilon})$ term is determined by considering a shift of the control variable in the normal direction and Taylor expanding it in the same direction.

$$\Phi(u^{\varepsilon}) = \Phi\,(\,u + \varepsilon\,\alpha\,n) = \Phi\,(u) + \varepsilon\,\alpha\,\frac{\partial \Phi(u)}{\partial n} + O(\varepsilon^2) \tag{5.5}$$

We can now evaluate the Equation (5.3) by using the relations derived in Equations (5.4) and (5.5).

$$\frac{\partial B}{\partial \theta} = \frac{1}{\varepsilon}[\,\nabla\Phi(u^{\varepsilon}).n^{\varepsilon} - \nabla\Phi(u).n\,]$$

$$= \frac{1}{\varepsilon}\left\{\nabla\,(\Phi + \varepsilon\,\alpha\,\frac{\partial \Phi}{\partial n}).(n - \varepsilon \sum_{t=\xi,\eta} \frac{\partial \alpha}{\partial t} t) \quad - \quad \nabla\,\Phi\,.\,n\right\}$$

$$= \frac{1}{\varepsilon}\left\{ \nabla\,\Phi . n - \varepsilon\,\nabla\,\Phi \sum_{t=\xi,\eta} \frac{\partial \alpha}{\partial t} t + \varepsilon\,\alpha\,\nabla(\frac{\partial \Phi}{\partial n}).n \;-\nabla\,\Phi\,.\,n + O(\varepsilon^2) \right\}$$

(5.6) $$\frac{\partial B}{\partial \theta} = -\sum_{t=\xi,\eta} \frac{\partial \alpha}{\partial t} t\,\nabla\Phi + \;\alpha\,(\frac{\partial^2 \Phi}{\partial n^2})$$

Let us now evaluate the term $\frac{\partial f}{\partial \theta}$. We know that f is an implicit function of the state (flow) variable. Perturbation of the flow variable due to the change of geometry is already given in the above Equation (5.5). Therefore, we here state the variation of the function f with respect to the geometry (control) variable;

$$\frac{\partial f}{\partial \theta} = \frac{1}{\varepsilon}\;\left\{\; f\,[\,\Phi\,(\,u^{\varepsilon}\,)\,] - f\,[\,\Phi\,(u)\,]\,\right\}$$

Making use of the Equation (5.5);

$$f\,[\,\Phi\,(\,u^{\varepsilon}\,)\,] = f\,[\,\Phi\,(u) + \varepsilon\,\alpha\,\frac{\partial \Phi(u)}{\partial n}\,]$$

Considering the second term in the brackets as the infinitesimal increment and by expanding f around the unperturbed geometry as given below, we can express the variation of the function on the original geometry of the body surface.

$$f\,[\,\Phi\,(u) + \varepsilon\,\alpha\,\frac{\partial \Phi(u)}{\partial n}\,] = f\,[\,\Phi\,(u)\,] \;+\; \varepsilon\,\alpha\,f_{\Phi}\,\frac{\partial \Phi}{\partial n} \;+\; O(\varepsilon^2)$$

Inserting this expression into the variational equation and by making use of the initial boundary condition, we can see that the term $\frac{\partial f}{\partial \theta}$ goes to zero if we consider only the change of the geometry by keeping the flow (state) variable fixed.

$$\frac{1}{\varepsilon}\;\left\{\; f\,[\,\Phi\,(\,u^{\varepsilon}\,)\,] - f\,[\,\Phi\,(u)\,]\,\right\} = \frac{1}{\varepsilon}\left\{ f\,[\,\Phi\,(u)\,] \;+\; \varepsilon\,\alpha\,f_{\Phi}\;\underbrace{\frac{\partial \Phi}{\partial n}}_{=0,\,BC.2}\;-\;f\,[\,\Phi\,(u)\,] \right\}$$

Therefore,

(5.7) $$\frac{\partial f}{\partial \theta} = 0$$

We can now bring the equations (5.2), (5.6) and (5.7) together to express the variational Equation (5.1).

$$\frac{\partial}{\partial \theta} \int_{BS} (f - \lambda B)\, ds = \int_{BS} (f - \lambda B) \frac{\partial}{\partial \theta}(ds) + \int_{BS} \left(\frac{\partial f}{\partial \theta} - \lambda \frac{\partial B}{\partial \theta}\right) ds$$

$$= \int_{BS} \left\{ \frac{\alpha}{R} f \;+\; \lambda \sum_{t=\xi,\eta} \frac{\partial \alpha}{\partial t} t\, \nabla\Phi \;-\; \alpha\, \lambda \left(\frac{\partial^2 \Phi}{\partial n^2}\right) \right\} ds$$

(5.8) $$\frac{\partial}{\partial \theta} \int_{BS} (f - \lambda B)\, ds = \int_{BS} \left\{ \frac{\alpha}{R} f \;+\; \lambda \;\left[\sum_{t=\xi,\eta} \frac{\partial \alpha}{\partial t} t\, \nabla\Phi \;-\; \alpha \left(\frac{\partial^2 \Phi}{\partial n^2}\right) \right] \right\} ds$$

Chapter 6

6.1 Variation of the Objective Functional with respect to the State Variable

This part of the study is devoted to the formulation to determine the variation of the objective functional with respect to the state variable where the objective functional is a function of the state. The formulation of this variation is important both for the extensive implementations in the optimal control theory and for the solution of the adjoint operators as we have given as an example in the previous chapter. Two alternative ways have been considered and they have proved to be giving identical results when compared to each other at the end.

We will first formulate the problem by expanding the perturbation in power series. This is a more general formulation compared to the second approach where we convert the objective functional from an implicit function of the state into an explicit form by restating it in terms of the velocities components. Both methods have been observed to give identical results.

6.2 First Approach : Expanding the perturbation in power series

We will consider the same domain that we have introduced in Figure 5-1. This fluid domain Ω is perturbed as a result of the perturbations on its boundary by changing the geometry of body surface BS to find its optimal shape for the stated objective functional. As we have previously stated the objective functional is given as

$$J(u,x) = \int_{BS} f(x)\, ds$$

where J is the objective functional of the control and the state variables determined by the surface integration over the body surface and f is the functional defined as an implicit functional of the state variable. We should here again state that based on the nature of the shape optimization problem, our control variable u is a variable θ that defines the shape of the body and the velocity potential ϕ is the variable that specifies the state of the flow; i.e. the state variable.

For a small perturbation ε of the objective functional, its variation is defined as;

$$\frac{1}{\varepsilon}[\, J(BS^{\varepsilon}, x^{\varepsilon}) - J(BS, x)\,]$$

In the above statement, the superscript ε denotes the perturbed form of the term, as in the perturbation of BS and the state (flow) variable x. The state variable is given as a function of the domain boundary to be optimized, i.e. the body surface.

$$x = x(BS)$$

The change in the incremental element ds for the surface integration should also be considered in this problem. However, as we will see later, this perturbation gives us an additional term and therefore, it will be considered as an additional part at the end of this formulation. For now, by neglecting the term ds^{ε}, we will exclude the small change in the length of the incremental surface element, yet the problem will be considered over the perturbed boundary of the domain; BS^{ε}.

In order to express the variation of the functional;

$$\frac{1}{\varepsilon}[\, \int_{BS^{\varepsilon}} f^{\varepsilon}(x^{\varepsilon})ds - \int_{BS} f(x)ds\,]$$

we should express the first term in the brackets in terms of an integral and quantities on the unperturbed surface BS.

Let us consider a geometry defined by the control variable u. A corresponding shift of the control variable is given by;

$$u^{\varepsilon} = u + \varepsilon\alpha\vec{n}$$

Here, the perturbation is defined in the normal direction. α is an arbitrary function defined on the boundary. State (flow) variable x is a function of the control variable, i.e. the geometry of the body to be optimized. If we consider the value of the state variable for the perturbed control given above by Taylor expanding it;

$$x(u^{\varepsilon}) = x(u + \varepsilon\alpha n) = x(u) + \varepsilon\alpha\frac{\partial x}{\partial n} + O(\varepsilon^2)$$

And applying the same perturbation to the state variable for the above perturbed control variable as;

$$x^{\varepsilon}(u^{\varepsilon}) = x(u + \varepsilon\alpha n) + \varepsilon\delta x(u + \varepsilon\alpha n) + O(\varepsilon^2)$$

where δx is a small variation of the state variable. We can one more time expand this expression in Taylor series and assuming a small perturbation, we consider only the terms of the order $O(\varepsilon)$;

$$x^{\varepsilon}(u^{\varepsilon}) = x(u) + \varepsilon\alpha\frac{\partial x(u)}{\partial n} + \varepsilon\delta x(u) + \underbrace{\varepsilon^2\alpha\delta x\frac{\partial x(u)}{\partial n} + O(\varepsilon^2)}_{O(\varepsilon^2)}$$

Ignoring the higher order terms in the above expression, we can now determine the value of the functional f, as a function of the perturbed state (flow) variable but evaluated only in terms of the unperturbed flow variable and unperturbed body geometry.

$$f(x^{\varepsilon})\Big|_{BS^{\varepsilon}} = f\left[x(u) + \varepsilon\alpha\frac{\partial x(u)}{\partial n} + \varepsilon\delta x(u)\right]\Bigg|_{BS} + O(\varepsilon^2)$$

Assuming for now that the functional f does not explicitly depend on the control variable u, we can reorganize the above terms by considering the last two of them as the increment of the Taylor expansion.

$$f(x^{\varepsilon})\Big|_{BS^{\varepsilon}} = f(x)\Big|_{BS} + \varepsilon\alpha\frac{\partial x}{\partial n} f_x(x)\Big|_{BS} + \varepsilon\delta x f_x(x)\Big|_{BS} + O(\varepsilon^2)$$

Later the area element is included and the variation of the functional is evaluated.

$$\frac{1}{\varepsilon}[\,J(BS^{\varepsilon},x^{\varepsilon}) - J(BS,x)\,] = \frac{1}{\varepsilon}[\,\int_{BS^{\varepsilon}} f^{\varepsilon}(x^{\varepsilon})ds - \int_{BS} f(x)ds\,] =$$

$$= f(x^{\varepsilon})\Big|_{BS^{\varepsilon}} - f(x)\Big|_{BS} = f(x)\Big|_{BS} + \varepsilon\alpha\frac{\partial x}{\partial n} f_x(x)\Big|_{BS} + \varepsilon\delta x f_x(x)\Big|_{BS} - f(x)\Big|_{BS}$$

$$= \frac{1}{\varepsilon}[\,\int_{BS^{\varepsilon}} f^{\varepsilon}(x^{\varepsilon})ds - \int_{BS} f(x)ds\,] = \frac{1}{\varepsilon}\varepsilon\int_{BS}\left\{\alpha\frac{\partial x}{\partial n} f_x(x) + \delta x f_x(x)\right\}ds$$

Finally, we obtain the following equation for the variation of the functional with respect to the state variable, where the functional is given as an implicit function of it.

(6.1)
$$\frac{1}{\varepsilon}[\,\int_{BS^{\varepsilon}} f^{\varepsilon}(x^{\varepsilon})ds - \int_{BS} f(x)ds\,] = \int_{BS}[\,\alpha\frac{\partial x}{\partial n} f_x(x) + \delta x f_x(x)\,]\,ds$$

Let us now consider a practical case to apply the above equation for an objective functional to determine the wave resistance as given below;

$$J = -\int_{BS} p\, n_x ds$$

where p is the pressure and n_x is the x-component of the unit normal. Let us take p as our state variable (x = p in Equation 5.1) and later determine the implicit dependence on the total velocity potential ϕ by invoking the Bernoulli equation defined on a shore-based coordinate system. So, based on the Bernoulli equation we define p as;

$$p = \frac{(p - p_{\infty})}{\rho} = -[\,\frac{\partial\phi}{\partial t} + \frac{1}{2}|\nabla\phi|^2 + gz\,]$$

Assuming that we maintain the average depth of the body, we can ignore the hydrostatic component above. For the simplicity of the relations, let us define another function q with its relation to the total velocity potential given below;

$$p = f(q), \qquad q = \frac{\partial \phi}{\partial t} + \frac{1}{2}\left|\nabla \phi\right|^2$$

As a result of changes in the geometry of the body surface BS, the total velocity potential changes to $\phi + \delta\phi + O(\varepsilon^2)$ and p changes to $p + \delta p + O(\varepsilon^2)$. Based on the above relations between p, q and ϕ, we can express δp in terms of ϕ *and* $\delta\phi$;

$$\delta p = \frac{\partial f}{\partial q}\left(\frac{\partial \delta\phi}{\partial t} + \nabla\phi . \nabla\delta\phi\right)$$

Let us consider a curvilinear coordinate system on the body surface with two tangential coordinates; ξ *and* η and a normal coordinate ζ. The above equation can then be expressed as;

$$\delta p = \frac{\partial f}{\partial q}\frac{\partial \delta\phi}{\partial t} + \frac{\partial f}{\partial q}\frac{\partial \phi}{\partial \zeta}\frac{\partial \delta\phi}{\partial \zeta} + \frac{\partial f}{\partial q}\frac{\partial \phi}{\partial \xi}\frac{\partial \delta\phi}{\partial \xi} + \frac{\partial f}{\partial q}\frac{\partial \phi}{\partial \eta}\frac{\partial \delta\phi}{\partial \eta}$$

A boundary condition on the body surface that imposes a no-penetration condition cancels the second term of the above equation with the normal coordinate of the body. Therefore, the equation reduces to

$$\delta p = \quad \frac{\partial f}{\partial q}\frac{\partial \delta\phi}{\partial t} + \frac{\partial f}{\partial q}\frac{\partial \phi}{\partial \xi}\frac{\partial \delta\phi}{\partial \xi} + \frac{\partial f}{\partial q}\frac{\partial \phi}{\partial \eta}\frac{\partial \delta\phi}{\partial \eta} = \quad \frac{\partial f}{\partial q}\frac{\partial \delta\phi}{\partial t} + \frac{\partial f}{\partial q}\nabla_T\phi . \nabla_T\delta\phi$$

Where ∇_T denotes the operation on the tangential coordinates; ξ *and* η only. Let us now go back to our general equation of variation (Eq.6.1) and substitute the relevant terms for our problem with the δp term already determined above.

$$f_x = -n_x\,, \qquad \frac{\partial x}{\partial n} = \frac{\partial p}{\partial \zeta}$$

$$\int_{BS} [\, \alpha \frac{\partial x}{\partial n} f_x(x) \;+\; \delta x f_x(x) \,] \, ds = \int_{BS} [(-n_x)\alpha \frac{\partial p}{\partial n} + (-n_x)\delta p] \, ds$$

Let us evaluate the term $\frac{\partial p}{\partial n}$ in the above expression;

$$\frac{\partial p}{\partial n} = \frac{\partial f}{\partial q}\frac{\partial q}{\partial n} = \frac{\partial f}{\partial q}\frac{\partial}{\partial n}[\;\; \frac{\partial \phi}{\partial t} + \frac{1}{2}\left|\nabla \phi\right|^2 \;\;] = \frac{\partial f}{\partial q}\frac{\partial}{\partial t}(\frac{\partial \phi}{\partial \zeta}) + \frac{1}{2}\frac{\partial f}{\partial q}\frac{\partial}{\partial n}(\left|\frac{\partial \phi}{\partial \zeta}\right|^2 + \left|\nabla_T \phi\right|^2)$$

Where we have changed the order of differentiation to make use of the body boundary condition in the first term and separated the second term into normal and tangential coordinates. Since tangential velocity terms have no component in the normal direction and the no-penetration body boundary condition applies for both of the terms on the right hand side of the equation, the term $\frac{\partial p}{\partial n}$ vanishes. Therefore, our equation reduces to

$$\int_{BS} [(-n_x)\alpha \frac{\partial p}{\partial n} + (-n_x)\delta p] \, ds = \int_{BS} [\, (-n_x)(\frac{\partial f}{\partial q}\frac{\partial \delta\phi}{\partial t} + \frac{\partial f}{\partial q}\nabla_T \phi . \nabla_T \delta\phi) \,] \, ds$$

Substituting $\frac{\partial f}{\partial q} = 1$, we can also apply integration by parts for the second term to leave the admissible variation of the total velocity potential $\delta\phi$ alone.

$$\delta J = \int_{BS} (-n_x)\frac{\partial \delta\phi}{\partial t} \, ds \;\; + \;\; [\, \delta\phi\,(-n_x)\,\nabla_T \phi \,]\,\big|_{BS} \; - \; \int_{BS} \delta\phi \left\{ \frac{\partial}{\partial \xi}[\,(-n_x)\frac{\partial \phi}{\partial \xi}\,] + \frac{\partial}{\partial \eta}[\,(-n_x)\frac{\partial \phi}{\partial \eta}\,] \right\} ds$$

The second term in the above equation cancels because the surface integration is independent of the path and the tangential coordinates of the body surface starts and ends at the same point for a closed body. Therefore we obtain the following equation for the variation of the objective functional due to the state variable where the objective functional is an implicit function of the state.

$$(6.2) \qquad \delta J = -\int_{BS} \delta\phi \left\{ \frac{\partial}{\partial \xi}[(-n_x)\frac{\partial \phi}{\partial \xi}] + \frac{\partial}{\partial \eta}[(-n_x)\frac{\partial \phi}{\partial \eta}] \right\} ds + \int_{BS} (-n_x)\frac{\partial \delta\phi}{\partial t} ds$$

The first term in the above equation is in fact what we need for the solution of the adjoint equation (2) in Chapter 5-1. The variation of the objective functional with respect to the state will also find its extensive use in the more general optimal control problems. The second term in the above equation seems to be a little more complicated to define which takes the unsteadiness into consideration. However, once the dynamic condition is determined for the optimal control problem, we will have an equation for $\dot{x} = \frac{\partial \phi}{\partial t}$ and an admissible variation of the velocity potential $\delta\phi$ will also satisfy the dynamic condition given as a function of x(t);state and u(t);control.

6.3 Contribution from the independent variation of control

We have formulated the variation of the objective functional without taking the change in the length of the incremental surface element ds into account. This independent contribution of the control variable appears as an additional term as we will quickly include in the following lines.

We define the incremental piece ds as an arc

$$ds = R\, d\theta$$

where R is the radius of curvature and $d\theta$ is the incremental angle seeing the piece of the body surface. A small change of geometry in the normal direction is given by

$$-\varepsilon\, \alpha\, \vec{n}$$

As we have defined previously, α is an arbitrary function defined on the boundary of the surface. Therefore the change in the incremental element due to the above defined change of the geometry in the normal direction is

$$R\,d\theta - \varepsilon\,\alpha\,d\theta = (R - \varepsilon\,\alpha)\,d\theta = R\,(1 - \varepsilon\,\frac{\alpha}{R})\,d\theta$$

By substituting ds in the above equation we get the change in the incremental element;

(6.3) $$ds^{\varepsilon} = (1 - \varepsilon\,\frac{\alpha}{R})\,ds$$

The radius of curvature R for a surface element is determined by taking the product of the two line elements of the tangential coordinates in the curvilinear coordinate system.

$$\frac{1}{R} = \frac{1}{R_{\xi}} + \frac{1}{R_{\eta}}$$

Equation 6.3 leads to an additional term in the calculation of the variational formulation. For a more accurate result the perturbed objective functional term $\int_{BS^{\varepsilon}} f^{\varepsilon}(x^{\varepsilon})ds$ is now multiplied with ds^{ε} instead of ds;

$$f(x^{\varepsilon})ds^{\varepsilon} = [\,f\,(x) + \varepsilon\alpha\frac{\partial x}{\partial n}f_x(x) + \varepsilon\delta x f_x(x)\,]\,(1 - \varepsilon\,\frac{\alpha}{R})\,ds$$

Because of the small change in the length on the order of ε, we ignore higher order terms and

$$f(x^{\varepsilon})ds^{\varepsilon} - f(x)ds = \varepsilon[\;\alpha\frac{\partial x}{\partial n}f_x(x) \;+\; \delta x f_x(x)\;]\,ds \;-\; \varepsilon\,\frac{\alpha}{R}f(x)ds$$

Finally, the last term on the right hand side of the above equation appears as an additional term to the variational Equation 6.1;

$$\frac{1}{\varepsilon}[\;\int_{BS^{\varepsilon}} f^{\varepsilon}(x^{\varepsilon})ds - \int_{BS} f(x)ds\,] = \int_{BS}[\,\alpha\frac{\partial x}{\partial n}f_x(x) + \delta x f_x(x)\,]\,ds - \int_{BS}\frac{\alpha}{R}f(x)ds$$

6.4 Second Approach: Variation defined in terms of Velocity Components

We will verify the formulation of the objective functional variation with respect to the state variable by expressing the implicit function in terms of velocity components.

Let us consider the same objective functional to determine the wave resistance as given below;

$$J = -\int_{BS} p\, n_x ds$$

where p is the pressure and n_x is the x-component of the unit normal. Again p will be our state variable (x = p) and Bernoulli equation will be revisited to express the implicit dependence of the state variable in terms of velocities. The problem is formulated on a fixed coordinate system. We define p as;

$$p = \frac{(p - p_\infty)}{\rho} = -[\frac{\partial \phi}{\partial t} + \frac{1}{2}|\nabla \phi|^2\ ]$$

Assuming that the total velocity potential is composed of a flow in the –x direction and a perturbation potential such that;

$$\Phi = -Ux + \phi$$

The $|\nabla\Phi|^2$ term is then calculated as;

$$|\nabla\Phi|^2 = (U + \frac{\partial \phi}{\partial x})^2 + (\frac{\partial \phi}{\partial y})^2 + (\frac{\partial \phi}{\partial z})^2$$

We can now substitute this into the expression for p;

$$p = -[\frac{\partial \phi}{\partial t} + \frac{1}{2}|\nabla \phi|^2\ ] = -[\frac{\partial \phi}{\partial t} + \frac{1}{2}\left\{(U + \frac{\partial \phi}{\partial x})^2 + (\frac{\partial \phi}{\partial y})^2 + (\frac{\partial \phi}{\partial z})^2\right\}]$$

$$= -\frac{\partial \phi}{\partial t} - \frac{1}{2}\left\{(\frac{\partial \phi}{\partial x})^2 + (\frac{\partial \phi}{\partial y})^2 + (\frac{\partial \phi}{\partial z})^2\right\} - \frac{1}{2}U^2 - U\frac{\partial \phi}{\partial x}$$

The objective function is a function of P. Therefore, its variation is given by

$$\delta J = \frac{\partial f}{\partial P}\delta P$$

However, P is a function of the state variable, in our case, ϕ. So, we are able to express this implicit variation in terms of the variation of the state variable. For this, we should take the variation of the above equation giving p in terms of state variable.

$$\partial p = -\frac{\partial \delta\phi}{\partial t} - \left\{\frac{\partial \phi}{\partial x}\frac{\partial \delta\phi}{\partial x} + \frac{\partial \phi}{\partial y}\frac{\partial \delta\phi}{\partial y} + \frac{\partial \phi}{\partial z}\frac{\partial \delta\phi}{\partial z}\right\} - U\frac{\partial \delta\phi}{\partial x}$$

by ordering the terms and denoting the perturbation velocities in the x, y and z directions as u, v and w respectively;

$$\partial p = -(U+u)\frac{\partial \delta\phi}{\partial x} - v\frac{\partial \delta\phi}{\partial y} - w\frac{\partial \delta\phi}{\partial z} - \frac{\partial \delta\phi}{\partial t}$$

In order to be able to make use of the no-penetration condition on the body surface, conversion to a curvilinear coordinate system would be beneficial. For this aim, we introduce the following coordinate conversion;

$$x = x(\xi,\eta,\zeta) \;\; ; \;\; y = y(\xi,\eta,\zeta) \;\; ; \;\; z = z(\xi,\eta,\zeta)$$

where again, ξ *and* η are the tangential coordinates on the body surface and ζ is the normal coordinate. Let us now substitute the curvilinear coordinates into the Cartesian coordinates given above.

$$\begin{aligned}\partial p = -\frac{\partial \delta\phi}{\partial t} - (U+u)&\left\{\frac{\partial \delta\phi}{\partial \xi}\xi_x + \frac{\partial \delta\phi}{\partial \eta}\eta_x + \frac{\partial \delta\phi}{\partial \varsigma}\varsigma_x\right\} \\ -v&\left\{\frac{\partial \delta\phi}{\partial y}\xi_y + \frac{\partial \delta\phi}{\partial \eta}\eta_y + \frac{\partial \delta\phi}{\partial \varsigma}\varsigma_y\right\} \\ -w&\left\{\frac{\partial \delta\phi}{\partial z}\xi_z + \frac{\partial \delta\phi}{\partial \eta}\eta_z + \frac{\partial \delta\phi}{\partial \varsigma}\varsigma_z\right\}\end{aligned}$$

Rearranging the terms above for bringing the terms with the same curvilinear coordinates together will enable us to apply the no-penetration condition easily by stating that the curvilinear velocity component in the ς (normal) direction is equal to zero.

$$\begin{aligned}\partial p = -\frac{\partial\delta\phi}{\partial t} - \frac{\partial\delta\phi}{\partial\xi}\ &\left\{(U+u)\,\xi_x + v\,\xi_y + w\,\xi_z\right\}\\ -\frac{\partial\delta\phi}{\partial\eta}\ &\left\{(U+u)\,\eta_x + v\,\eta_y + w\,\eta_z\right\}\\ \underbrace{-\frac{\partial\delta\phi}{\partial\varsigma}\ \left\{(U+u)\,\varsigma_x + v\,\varsigma_y + w\,\varsigma_z\right\}}_{=0,\ no-penetration\ condition}\end{aligned}$$

Let us denote the tangential velocity components given in the first and the second lines of the above expression with E *and* N respectively and restate the variation.

$$\partial p = -\frac{\partial\delta\phi}{\partial t} - \frac{\partial\delta\phi}{\partial\xi}\mathrm{E} - \frac{\partial\delta\phi}{\partial\eta}\mathrm{N}$$

The variation of the objective functional is then given again;

$$\delta J = \int_{BS}\frac{\partial f}{\partial p}\,\delta p\,ds = -\int_{BS}(-n_x)\left\{\frac{\partial\delta\phi}{\partial\xi}\mathrm{E} + \frac{\partial\delta\phi}{\partial\eta}\mathrm{N}\right\}ds + \int_{BS}(-n_x)\frac{\partial\delta\phi}{\partial t}\,ds$$

We can apply integration by parts to leave the admissible variation of the state alone, as we did previously

$$= -[\,\delta\phi\,(-n_x)\,(\mathrm{E}+\mathrm{N})\,]\,\big|_{BS} - \int_{BS}\delta\phi\left\{\frac{\partial}{\partial\xi}[\,(-n_x)\,\mathrm{E}\,] + \frac{\partial}{\partial\eta}[\,(-n_x)\,\mathrm{N}\,]\right\}ds + \int_{BS}(-n_x)\frac{\partial\delta\phi}{\partial t}\,ds$$

Due to the geometry of the closed body, the first term vanishes in the above integration and we obtain the following result for the variation of the objective functional with respect to the state variable.

$$\delta J = -\int_{BS}\delta\phi\left\{\frac{\partial}{\partial\xi}[\,(-n_x)\,\mathrm{E}\,] + \frac{\partial}{\partial\eta}[\,(-n_x)\,\mathrm{N}\,]\right\}ds + \int_{BS}(-n_x)\frac{\partial\delta\phi}{\partial t}\,ds \qquad (6.4)$$

With E and N being the tangential velocity components in the curvilinear coordinates of ξ *and* η respectively, we can see that this approach gives us the same equation as Equation (6.2) that we have derived previously.

The second approach does not consider that the perturbation of the velocity potential is in fact caused by the perturbation on the geometry of the body, i.e. the state variable is a function of the control variable. Therefore, the additional term taking the change in the length of the incremental element into account (given in Equation 6.3) is not considered in this derivation. However, when we take the gradient of the objective functional with respect to the control variable for the formulation of the optimal control theory, or as we did previously for the simplistic approach to the problem by means of taking the reduced gradient, this contribution is included, too.

Chapter 7

7.1 Application of the Lagrange Multiplier Theorem to the Shape Optimization Problem

In Chapter 5.1 we have defined a problem where a three dimensional body is translating with a steady forward velocity U well below a free surface. We have formulated the problem in terms of the optimal control theory. Our aim has been to optimize the shape of the body for a general objective function which is a function of both the state and the control variable.

In this part of the study, we would like to apply the Lagrange multiplier theorem briefly introduced as in Chapter-3 to the same optimization problem directly to compare with the results of the formulation based on the optimal control theory.

Although both of the methods are benefiting from the convenience of the general Lagrange Multiplier Theorem and the adjoining of the constraint for finding an optimum solution, there exists a difference in the way followed. It arises from the fact that optimal control approach implements the stationary condition of the Lagrangian by disconnecting the dependence of the state and control and treating them as independent variables. In this approach, we have defined our control (u) as the perturbation of the geometry, denoted with θ (a set of geometric parameters to define the body surface which is to be optimized) and for the convenience of defining the state (x) of the flow, we have chosen the velocity potential ϕ as our state variable.

However, in order to define the same problem in terms of the Lagrange Multiplier Theorem given in Chapter-3, we express the complete dependence of the variables without disconnecting their relations which therefore causes a little more complexity and difficulty in formulization. In other words, for optimizing an objective functional which is an implicit function of the flow (state) variable, we take the objective functional as a function of the flow (state) variable, and at the same time, the state variable is the function of the control (geometry) variable u.

Let us now start with the formulation of the problem. The sketch describing the flow domain to be considered given in Chapter-5.1 is also valid for this formulation. In order first to compare the results of the two approaches, we will locate the three dimensional body well under a free surface. Later in this study, we will consider the addition of the free surface which has a more physical meaning from the practical point of view.

The constraint space H (u) is composed of a set of boundary conditions which maps the design variables from the primal space X into the constraint space Z.

(B.C.1) $$\nabla^2 \phi = 0 \qquad (\text{in}\,\Omega)$$

(B.C.2) $$\frac{\partial \phi}{\partial n} + \vec{U}.\vec{n} = 0 \qquad (\text{on BS})$$

(B.C.3) $$\phi = 0 \qquad (\text{on FFS})$$

Note that we will need to express the flow variable (x) as a function of the design variable u. And our design variable is a function defined on the surface of the body. In order to be able to invoke the continuity condition which is valid throughout the whole fluid domain, we will benefit from Green's theorem to distribute it on the boundaries of the fluid domain. Since we perturb the flow variable only by means of the control variable without disconnecting their dependence, we will later choose appropriate Lagrange Multipliers so that the variation of the flow itself will vanish. Lagrange Multipliers play a very important role here on each boundary of the fluid domain for determining the relation of the flow and control variables such that the flow perturbation is caused only by the variation of the control variable, i.e. by means of the perturbation of the body surface.

We consider a general objective function in the form of

$$J(u,x) = \int_{BS} f \, ds = \int_{BS} p \, n_x \, ds$$

for minimizing the wave resistance of the submerged body. In Chapter 6-1, we have already determined the variation of the above objective function with respect to the flow variable, where the variation is caused both by the change of geometry and the variation of the flow variable itself. The relevant equation is given here again:

$$\frac{1}{\varepsilon}\left[\int_{BS^{\varepsilon}} f^{\varepsilon}(x^{\varepsilon})ds - \int_{BS} f(x)ds \right] = \int_{BS} \left[\alpha \frac{\partial x}{\partial n} f_x(x) + \delta x f_x(x) \right] ds - \int_{BS} \frac{\alpha}{R} f(x)ds$$

In the same Chapter, we have also already applied the above equation to our above defined objective functional to minimize the wave resistance. It is given in Equation (6.2) and with the inclusion of the change in the incremental surface element given in part 6.2, we restate the variation of the objective functional below.

$$(7.1) \quad \delta J = - \int_{BS} \delta\phi \left\{ \frac{\partial}{\partial \xi}\left[(-n_x)\frac{\partial \phi}{\partial \xi} \right] + \frac{\partial}{\partial \eta}\left[(-n_x)\frac{\partial \phi}{\partial \eta} \right] \right\} ds - \int_{BS} \alpha \frac{p \, n_x}{R} ds + \int_{BS} (-n_x)\frac{\partial \delta\phi}{\partial t} ds$$

For now, we will neglect the last term of the above equation for simplicity. Note that the above equation includes the variation of the flow variable $\delta\phi$ that is independent of the body surface perturbation. Reviewing the formulization of the variation in Chapter-6 will make the above statement clear since as we may remember, we have included the perturbation of the flow such that

$$x^{\varepsilon}(u^{\varepsilon}) = x(u + \varepsilon\alpha n) + \varepsilon\delta x(u + \varepsilon\alpha n) + O(\varepsilon^2)$$

where x is the state (flow) variable and the equation included such terms as $\delta x \, (= \delta\phi)$. By means of a proper choice of the Lagrange Multipliers on each boundary of the fluid domain, the necessary conditions should be satisfied for an arbitrary admissible variation of the state $\delta x\ (=\delta\phi)$.

Let us now speak in terms of the Lagrange Multiplier Theorem. We have previously stated the necessary conditions of the local theory of optimization for an extremum of an objective functional f that is subject to the constraint $H(u)=\theta$ (i.e. null vector), where f is a real valued functional on a Banach space X (primal space) and H is a mapping from X into a Banach space Z (constraint space).

According to these conditions, if f and H are continuously Frechet differentiable in an open set containing the regular point u_0, and if f is assumed to achieve a local extremum that is subject to $H(u)=\theta$ at the point u_0; then $f'(u_0)h=0$ for all h satisfying $H'(u_0)h=\theta$.

Therefore, for any admissible variation of the state $\delta\phi$, at the local extremum u_0, this linear operator equates the Frechet differential of the transformation H (constraint equation in our problem) to null vector (=0 in our case, as the transformation H is simply a real-valued functional). As we know, the admissible variation $\delta\phi$ is then called the null space of the Frechet derivative of the constraint equation $H'(u_0)$. Following the necessary conditions above, again, the admissible variation $\delta\phi$ that is satisfying the above condition, also makes the Frechet differential of the real-valued objective functional f equal to zero.

Since both the objective functional and the constraint equation are real-valued, their combination forms the Lagrangian functional

$$L(u)=f(u)+z_0 * H(u)$$

Based on the above statements and the definition of the adjoint operators, this Lagrangian functional is stationary at the optimal point u_0, i.e.;

$$f'(u_0) + z_0 * H'(u_0) = \theta.$$

Just like the constraint equation $H(u_0)$, the linearized version of the constraint equation $H'(u_0)$ also maps the point u_0 from primal space X onto the constraint space Z with the regularity definition. The composition of adjoint of this linear operator, $H'(u_0)*$, with the Lagrange multiplier $z_0* \in Z*$ maps the optimization problem onto the dual of the primal space, X*. This is the normed Banach space where we have established the extension form of the projection theorem for our problem. From the application point of view, this approach is believed to provide to us a convenience to relate the null space and range of a linear operator and its adjoint.

Although in almost most of the optimization problems Lagrange multipliers are often treated as a convenient set of constants multiplied with the constraint equations, they are an entire set of functions defined in an appropriate dual space for each boundary of the domain. And they will be selected such that the stationarity condition of the Lagrangian will be satisfied at the optimal solution.

Let us now go back to our problem. Now that we have given the variation of the objective functional, it is to be combined with the constraint with a proper choice of the Lagrange multiplier on each boundary so that we can relate the constraint space and the dual of the primal space and satisfy the stationarity condition of the Lagrangian equation.

The constraint equation of continuity is valid throughout the whole fluid domain and its variation with respect to the flow variable is given by;

$$\nabla^2 \delta\phi = 0$$

It will be combined with a linear operator λ and will need to be distributed onto the boundaries of the fluid domain so that at the end, we will be able to choose the related control variable that is defined on the boundary of the body surface.

$$z_0 * H'(u) = \int_{\Omega} \lambda \nabla^2 \delta\phi \, d\Omega$$

Since the problem is given in a so-called 'discrete' domain that is composed of certain boundaries and the domain surrounded by them, we will make use of Green's theorem to express the above constraint in such a sense.

$$\int_{\Omega} \lambda \nabla^2 \delta\phi \, d\Omega - \int_{\Omega} \delta\phi \nabla^2 \lambda \, d\Omega = \int_{\substack{boundaries \\ of\, the\, fluid\, domain}} \left(\lambda \frac{\partial \delta\phi}{\partial n} - \delta\phi \frac{\partial \lambda}{\partial n} \right) ds$$

Leaving the first term on the left hand side of the equation alone,

$$z_0 * H'(u) = \int_{\Omega} \lambda \nabla^2 \delta\phi \, d\Omega \;=\; \int_{\Omega} \delta\phi \nabla^2 \lambda \, d\Omega + \int_{\substack{boundaries \\ of\, the\, fluid\, domain}} \left(\lambda \frac{\partial \delta\phi}{\partial n} - \delta\phi \frac{\partial \lambda}{\partial n} \right) ds$$

Let us express the boundaries of the fluid domain individually to find the corresponding Lagrange multipliers.

$$(7.2) \qquad z_0 * H'(u) = \int_{\Omega} \delta\phi \nabla^2 \lambda \, d\Omega + \int_{BS} \left(\lambda \frac{\partial \delta\phi}{\partial n} - \delta\phi \frac{\partial \lambda}{\partial n} \right) ds + \int_{FFS} \left(\lambda \frac{\partial \delta\phi}{\partial n} - \delta\phi \frac{\partial \lambda}{\partial n} \right) ds$$

In order to be able to express the body surface integration above over the unperturbed geometry, we need to transfer the no-penetration condition of the perturbation in terms of and as an integration of the original geometry. The incremental variation will still satisfy the no-penetration condition given with

$$\nabla \Phi^{\varepsilon}(u^{\varepsilon}).n^{\varepsilon}\Big|_{BS^{\varepsilon}} = 0$$

In Chapter-5.1, we have derived the perturbation of the flow variable due to the perturbation of the geometry as given in Equation (5.5) and also the change in the normal vector as given in Equation (5.4). What we need to do then is to include the perturbation of the flow variable itself by following the same expansion and find the product of it with the perturbed normal vector.

(5.5) $$\Phi(u^{\varepsilon}) = \Phi\,(\,u + \varepsilon\,\alpha\,n)$$

$$\Phi^{\varepsilon}(u^{\varepsilon}) = \Phi\,(\,u + \varepsilon\,\alpha\,n) + \varepsilon\,\delta\phi\,(u + \varepsilon\,\alpha\,n) + O(\varepsilon^{2})$$

$$\phi^{\varepsilon}(u^{\varepsilon}) = \phi(\,u) + \varepsilon\,\alpha\,\frac{\partial\phi(\,u)}{\partial n} + \varepsilon\,\delta\phi\,(u) + O(\varepsilon^{2})$$

(5.4) $$n^{\varepsilon} = \vec{n} - \varepsilon\sum_{t=\xi,\eta}\frac{\partial\alpha}{\partial t}\vec{t}$$

$$\nabla\Phi^{\varepsilon}(u^{\varepsilon}).n^{\varepsilon}\Big|_{BS^{\varepsilon}} = \nabla\,[\,\phi(\,u) + \varepsilon\,\alpha\,\frac{\partial\phi(\,u)}{\partial n} + \varepsilon\,\delta\phi\,(u)\,]\Bigg|_{BS}.[\,\vec{n} - \varepsilon\sum_{t=\xi,\eta}\frac{\partial\alpha}{\partial t}\vec{t}\,]\Bigg|_{BS} = 0$$

$$\nabla\Phi^{\varepsilon}(u^{\varepsilon}).n^{\varepsilon}\Big|_{BS^{\varepsilon}} = \underbrace{\nabla\phi.\,n}_{=0} - \varepsilon\,\nabla\phi\sum_{t=\xi,\eta}\frac{\partial\alpha}{\partial t}\vec{t} + \varepsilon\,\nabla\delta\phi.\,\vec{n} + \varepsilon\,\alpha\,\nabla(\frac{\partial\phi}{\partial n}).n + O(\varepsilon^{2}) = 0$$

$$\varepsilon\;[\,\nabla\delta\phi.\,\vec{n}\; + \alpha\,(\frac{\partial^{2}\phi}{\partial n^{2}}) - \nabla\phi\sum_{t=\xi,\eta}\frac{\partial\alpha}{\partial t}\vec{t}\;] = 0$$

Therefore, we can find an expression for the no-penetration condition satisfied by the incremental variation in terms of the unperturbed geometry as given below;

(7.3) $$\nabla\delta\phi.\,\vec{n}\,\Big|_{BS} = -\,\alpha\,(\frac{\partial^{2}\phi}{\partial n^{2}}) + \nabla\phi\sum_{t=\xi,\eta}\frac{\partial\alpha}{\partial t}\vec{t}$$

We can now combine the Frechet differentials of the objective function (Equation 7.1) and the constraint equation along with its adjoint multiplier (Equation 7.2) which gives the variation of the Lagrangian and which according to the necessary conditions defined previously will be stationary once the optimal body shape is obtained.

$$f'(u) + z_0 * H'(u) = -\int_{BS} \delta\phi \left\{ \frac{\partial}{\partial\xi}[(-n_x)\frac{\partial\phi}{\partial\xi}] + \frac{\partial}{\partial\eta}[(-n_x)\frac{\partial\phi}{\partial\eta}] \right\} ds - \int_{BS} \alpha \frac{p\, n_x}{R}\, ds$$

$$+ \int_{\Omega} \delta\phi \nabla^2 \lambda \, d\Omega + \int_{FFS} \left(\lambda \frac{\partial\delta\phi}{\partial n} - \delta\phi \frac{\partial\lambda}{\partial n} \right) ds + \int_{BS} \left(\lambda \underbrace{\frac{\partial\delta\phi}{\partial n}}_{Equ.(6.3)} - \delta\phi \frac{\partial\lambda}{\partial n} \right) ds$$

Since the term $\frac{\partial\delta\phi}{\partial n}$ is given as an integration over the unperturbed geometry in the above equation, we will express it by means of the Equation (7.3) that we have derived above.

Let us reorganize the above terms with the incremental variation $\delta\phi$ so that we can exclude the dependence of the variation on this term by the proper choice of Lagrange multipliers. In other words to be able to find the optimal shape, we will determine an admissible variation such that the Lagrangian equation will go to zero as we approach the local extremum. Reorganizing and collecting the terms under the same integration sign,

$$f'(u) + z_0 * H'(u) = -\int_{BS} \delta\phi \left\{ \left(\frac{\partial}{\partial\xi}[(-n_x)\frac{\partial\phi}{\partial\xi}] + \frac{\partial}{\partial\eta}[(-n_x)\frac{\partial\phi}{\partial\eta}] \right) + \frac{\partial\lambda}{\partial n} \right\} ds$$

(7.4)

$$+ \int_{\Omega} \delta\phi \nabla^2 \lambda \, d\Omega + \int_{FFS} \left(\lambda \frac{\partial\delta\phi}{\partial n} - \delta\phi \frac{\partial\lambda}{\partial n} \right) ds$$

$$- \int_{BS} \left\{ \alpha \frac{p\, n_x}{R} + \lambda \left(\alpha \left(\frac{\partial^2\phi}{\partial n^2}\right) - \sum_{t=\xi,\eta} \frac{\partial\alpha}{\partial t} \vec{t}\, \nabla\phi \right) \right\} ds$$

We can now choose the adjoint multipliers on each boundary to allow the above equation to satisfy the necessary conditions for an admissible variation $\delta\phi$.

ADJOINT EQUATIONS:

(1) $$\frac{\partial \lambda}{\partial n} = -\left\{ \frac{\partial}{\partial \xi}[(-n_x)\frac{\partial \phi}{\partial \xi}] + \frac{\partial}{\partial \eta}[(-n_x)\frac{\partial \phi}{\partial \eta}] \right\} \qquad \text{(on BS)}$$

(2) $$\nabla^2 \lambda = 0 \qquad (\text{in } \Omega)$$

(3) $$\lambda = 0 \qquad \text{(on FFS)}$$

Comparing with the adjoint equations that we have found by means of the theory based on optimal control in Chapter 5-1, one can see that the second and the third equations are identical. We have also formulated the variation of the objective functional with respect to the flow variable in Chapter 6. Equations (6.2) and (6.4) both reveal the fact that adjoint equation (1) is also identical with the previously determined one.

(6.2) $$\frac{\partial \lambda}{\partial n} = -\frac{\partial f}{\partial \phi} = -\left\{ \frac{\partial}{\partial \xi}[(-n_x)\frac{\partial \phi}{\partial \xi}] + \frac{\partial}{\partial \eta}[(-n_x)\frac{\partial \phi}{\partial \eta}] \right\}$$

With the adjoint equations determined above, what remains from Equation (7.4) is an integral over the body surface.

(7.5) $$f'(u) + z_0 * H'(u) = -\int_{BS} \left\{ \alpha \frac{p\, n_x}{R} + \lambda \left(\alpha \left(\frac{\partial^2 \phi}{\partial n^2}\right) - \sum_{t=\xi,\eta} \frac{\partial \alpha}{\partial t} \vec{t}\, \nabla \phi \right) \right\} ds$$

The above equation will therefore perturb the body geometry in the optimal direction until a local extremum is obtained in which case the above Lagrangian will be equal to zero. It is interesting to see that the above formulation has given the identical result (Equation 5.13) with the optimal control theory formulization given in Chapter 5.3. Therefore, we can make sure the validity of both of the approaches for finding the optimal solution to the shape optimization problem.

Once we have made sure that the optimization routine works fine, the next step will be to include a free surface boundary in the same problem so as to consider the submergence of the body just near the free surface.

7.2 Addition of the Free Surface Boundary to the Problem

Let us now consider a three dimensional body which is translating with a steady forward velocity U just below a free surface. We locate the body at a shallower depth so that the free surface boundary condition is now included in the problem. The problem is defined on a translating coordinate system. Since the body is translating with a steady forward speed, a steady flow is defined in the opposite direction of the translation far upstream of the body with a velocity of U. The flow domain Ω is bounded by the body surface (BS), free surface (FS) and the far-field surface (FFS). The following diagram gives the domain considered in this problem.

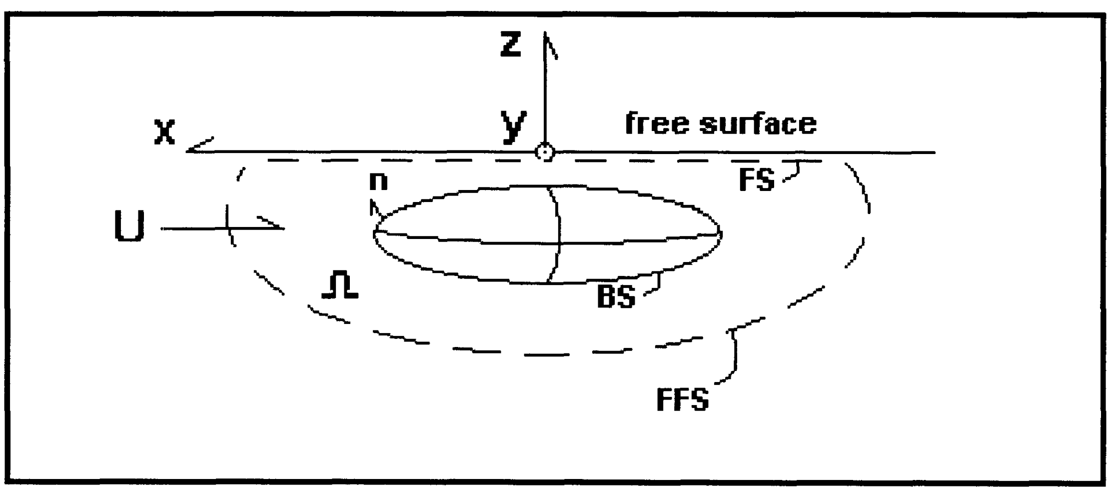

Figure-7-1: Sketch of the flow domain including the free surface boundary.

The constraint space H (u) now includes one additional boundary condition, given as (B.C.4) below which is the Neumann-Kelvin free surface condition governing the linear steady wave pattern generated by the translating body. We assume here, for the simplicity of the problem, that no ambient waves are present.

(B.C.1) $$\nabla^2 \phi = 0 \qquad (\text{in}\,\Omega)$$

(B.C.2) $$\frac{\partial \phi}{\partial n} + \vec{U}.\vec{n} = 0 \qquad (\text{on BS})$$

(B.C.3) $$\phi = 0 \qquad (\text{on FFS})$$

(B.C.4) $$U^2 \frac{\partial^2 \phi}{\partial x} + g \frac{\partial \phi}{\partial z} = \frac{\partial^2 \phi}{\partial x} + k \frac{\partial \phi}{\partial z} = 0 \;;\; (k = \frac{g}{U^2}) \qquad (\text{on FS})$$

The constraint equation of continuity is valid throughout the whole fluid domain and its variation with respect to the flow variable given by;

$$\nabla^2 \delta\phi = 0$$

will be combined with a linear operator λ and will need to be distributed onto the boundaries of the fluid domain so that at the end, we will be able to choose the related control variable that is defined on the boundary of the body surface.

$$z_0 * H'(u) = \int_{\Omega} \lambda \, \nabla^2 \delta\phi \, d\Omega$$

We will again make use of Green's theorem to express the above constraint in such a sense.

$$\int_{\Omega} \lambda \nabla^2 \delta\phi \, d\Omega - \int_{\Omega} \delta\phi \nabla^2 \lambda \, d\Omega = \int_{\substack{boundaries \\ of\ the\ fluid\ domain}} \left(\lambda \frac{\partial \delta\phi}{\partial n} - \delta\phi \frac{\partial \lambda}{\partial n} \right) ds$$

Leaving the first term on the left hand side of the equation alone,

$$z_0 * H'(u) = \int_\Omega \lambda \nabla^2 \delta\phi \, d\Omega \;=\; \int_\Omega \delta\phi \nabla^2 \lambda \, d\Omega + \int_{\substack{boundaries \\ of\ the\ fluid\ domain}} \left(\lambda \frac{\partial \delta\phi}{\partial n} - \delta\phi \frac{\partial \lambda}{\partial n} \right) ds$$

We will now have an additional boundary of the fluid domain which defines the free surface and this will give us an additional adjoint equation plus two adjoint boundary conditions to be satisfied at the end.

$$z_0 * H'(u) = \int_\Omega \delta\phi \nabla^2 \lambda \, d\Omega + \int_{BS} \left(\lambda \frac{\partial \delta\phi}{\partial n} - \delta\phi \frac{\partial \lambda}{\partial n} \right) ds + \int_{FS} \left(\lambda \frac{\partial \delta\phi}{\partial n} - \delta\phi \frac{\partial \lambda}{\partial n} \right) ds + \int_{FFS} \left(\lambda \frac{\partial \delta\phi}{\partial n} - \delta\phi \frac{\partial \lambda}{\partial n} \right) ds$$

We can now combine the Frechet differential of the objective function (Equation 7.1) and the above equation which gives the variation of the Lagrangian and which according to the necessary conditions defined previously will be stationary once the optimal body shape is obtained.

$$f'(u) + z_0 * H'(u) = -\int_{BS} \delta\phi \left\{ \frac{\partial}{\partial \xi} [\, (-n_x) \frac{\partial \phi}{\partial \xi} \,] + \frac{\partial}{\partial \eta} [\, (-n_x) \frac{\partial \phi}{\partial \eta} \,] \right\} ds - \int_{BS} \alpha \frac{p\, n_x}{R} \, ds$$

$$\int_\Omega \delta\phi \nabla^2 \lambda \, d\Omega + \int_{FS} \left(\lambda \frac{\partial \delta\phi}{\partial n} - \delta\phi \frac{\partial \lambda}{\partial n} \right) ds + \int_{FFS} \left(\lambda \frac{\partial \delta\phi}{\partial n} - \delta\phi \frac{\partial \lambda}{\partial n} \right) ds + \int_{BS} \left(\lambda \underbrace{\frac{\partial \delta\phi}{\partial n}}_{Equ.(6.3)} - \delta\phi \frac{\partial \lambda}{\partial n} \right) ds$$

We have already reorganized the above terms, except for the newly introduced free surface term, and have chosen the proper adjoint equations so that the necessary conditions are satisfied for any admissible variation $\delta\phi$ in the state variable. While doing this, we have enforced the no-penetration body boundary condition as defined in terms of the unperturbed geometry in Equation (7.3) and replaced the term $\frac{\partial \delta\phi}{\partial n}$ with it.

The same approached will be followed for the free surface term. We will enforce the free surface boundary condition (B.C.4) by noting that an admissible variation in the state variable $\delta\phi$ still complies with it.

$$\frac{\partial^2 \delta\phi}{\partial x^2} + k\frac{\partial \delta\phi}{\partial z} = 0 \qquad \text{(on FS)}$$

Without losing generality, we can assume that earth gravitational acceleration is large enough that the restoring role it plays leads to small wave slopes in general. This will enable us to linearize the free surface term that is included in the variation of the Lagrangian equation so that;

$$\int_{FS}\left(\lambda\frac{\partial \delta\phi}{\partial n} - \delta\phi\frac{\partial \lambda}{\partial n}\right)ds = \int_{FS}\left(\lambda\frac{\partial \delta\phi}{\partial z} - \delta\phi\frac{\partial \lambda}{\partial z}\right)ds$$

It is now possible to include the free surface boundary condition in the above equation by replacing the $\frac{\partial \delta\phi}{\partial z}$ term from (B.C.4).

$$\int_{FS}\left(\lambda\frac{\partial \delta\phi}{\partial z} - \delta\phi\frac{\partial \lambda}{\partial z}\right)ds = \int_{FS}\left(\lambda(-\frac{1}{k}\frac{\partial^2 \delta\phi}{\partial x}) - \delta\phi\frac{\partial \lambda}{\partial z}\right)ds$$

In order to leave the admissible variation of the state variable $\delta\phi$ alone which will enable us to define the proper adjoint equation, we will need to integrate the first term by parts twice. We define the limits of the free surface as [x_{up} , x_{down} , $-y$, y] to break the surface integration in terms of the coordinates x and y, as needed by the integration by parts. Applying it twice on the first term of the above equation;

$$-\frac{1}{k}\int_{FS}(\lambda\frac{\partial^2 \delta\phi}{\partial x})ds = \int_{-y}^{y}dy\int_{x_{down}}^{x_{up}}\lambda\frac{\partial^2 \delta\phi}{\partial x}dx = \int_{-y}^{y}dy\left[\ \lambda\frac{\partial \delta\phi}{\partial x}\bigg|_{x_{down}}^{x_{up}} - \int_{x_{down}}^{x_{up}}\frac{\partial \lambda}{\partial x}\frac{\partial \delta\phi}{\partial x}dx\ \right]$$

$$-\frac{1}{k}\int_{FS}(\lambda\frac{\partial^2 \delta\phi}{\partial x})ds = \int_{-y}^{y}dy\left[\ \lambda\frac{\partial \delta\phi}{\partial x} - \delta\phi\frac{\partial \lambda}{\partial x}\ \right]\bigg|_{x_{down}}^{x_{up}} + \int_{-y}^{y}dy\int_{x_{down}}^{x_{up}}\delta\phi\frac{\partial^2 \lambda}{\partial x^2}dx$$

by combining the above equation with the first term of the free surface integration;

$$\int_{FS}\left((-\frac{1}{k}\lambda\frac{\partial^2\delta\phi}{\partial x}) - \delta\phi\frac{\partial\lambda}{\partial z}\right)ds = \int_{FS}(-\frac{1}{k}\delta\phi\frac{\partial^2\lambda}{\partial x^2} - \delta\phi\frac{\partial\lambda}{\partial z})ds - \frac{1}{k}\int_{-y}^{y}dy\,[\ \lambda\frac{\partial\delta\phi}{\partial x} - \delta\phi\frac{\partial\lambda}{\partial x}\ ]\Bigg|_{x_{down}}^{x_{up}}$$

We can now bring the terms with the common factor $\delta\phi$ together and choose the appropriate adjoint equations.

$$\int_{FS}\left((-\frac{1}{k}\lambda\frac{\partial^2\delta\phi}{\partial x}) - \delta\phi\frac{\partial\lambda}{\partial z}\right)ds = \int_{FS}[-\delta\phi\,(\frac{1}{k}\frac{\partial^2\lambda}{\partial x^2} + \frac{\partial\lambda}{\partial z})\,]\,ds - \frac{1}{k}\int_{-y}^{y}dy\,[\ \lambda\frac{\partial\delta\phi}{\partial x} - \delta\phi\frac{\partial\lambda}{\partial x}\ ]\Bigg|_{x_{down}}^{x_{up}}$$

Based on the first term in the above equation, we can define the following adjoint equation on free surface boundary satisfied for any admissible variation of the state $\delta\phi$, in addition to the previously defined equations on the other boundaries of the domain.

(4) $$\frac{\partial^2\lambda}{\partial x^2} + k\frac{\partial\lambda}{\partial z} = 0 \qquad \text{(on FS)}$$

In addition to the adjoint equation defined above for the free surface boundary, we need to eliminate the boundary term given as the second term in the above equation;

$$-\frac{1}{k}\int_{-y}^{y}dy\,[\ \lambda\frac{\partial\delta\phi}{\partial x} - \delta\phi\frac{\partial\lambda}{\partial x}\ ]\Bigg|_{x_{down}}^{x_{up}}$$

For this, we need to define the radiation or Sommerfeld condition [16] at the far upstream boundary of the domain. The radiation condition as given below;

$$\phi,\ \frac{\partial\phi}{\partial x} \rightarrow 0 \qquad \text{(at } x_{up}\text{)}$$

is intended to preclude the possibility of an incoming wave generated at the upstream boundary of the domain. This is given as a necessary condition for the uniqueness of the solution. Therefore, the remainder of the above boundary term enables us to define the proper adjoint boundary conditions given as;

$$\lambda,\ \frac{\partial \lambda}{\partial x} \to 0 \qquad \text{(at } x_{down}\text{)}$$

With the above definition of the adjoint equation and the adjoint boundary conditions, free surface boundary is included in the optimization problem of the submerged body. We should note, before closing our discussion that since the adjoint equation of the free surface condition is defined on the free surface, it does not affect the solution of the submerged body problem. It would be in effect in the case where we have an objective function defined as a free surface integration (i.e. wave resistance calculation by means of a surface cut). Similarly, an optimization approach for a surface piercing body would as well necessitate one to solve the above free surface adjoint equation, since the perturbation of that part of the body geometry would also need to satisfy explicity the free surface condition, or the adjoint equivalent of it.

Chapter 8

Conclusions

A continuous shape optimization problem is considered in this study with the objective of reducing the wave resistance of a submerged body translating at a steady forward velocity well below a free surface.

In order to evaluate the gradient of the objective functional efficiently and accurately, the problem is formulated with an adjoint approach where the fluid domain governing equations are treated as constraints on the variations in flow variables. Two related approaches are considered in the formulation of the problem: The Lagrange multiplier theorem and optimal control theory.

In the optimal control theory formulation, when the constraint formed by the boundary conditions and the Laplace's governing equation is adjoined to the objective functional to construct the Lagrangian, the dependency of the state on the control is disconnected and they are treated as independent variables. The Lagrange multiplier has related these two variables with each other. Dependencies are preserved for the application of the Lagrange multiplier theorem which looks like a more straightforward formulation but is harder to solve. Both methods yielded identical solutions and adjoint equations to the problem.

Later, two alternative ways are considered for determining the variation of the objective functional with respect to the state variable. It is required to solve the adjoint equation defined on the body boundary. Comparison of these two ways also revealed identical solutions.

Finally, a free surface boundary is included in the optimization problem. Its effect on the submerged body shape optimization problem is considered.

The continuous nature of the shape optimization problem at hand and the relatively easier case governed by Laplace's equation has provided us with the flexibility to consider the problem in terms of local optimization by simply necessitating the differentiability of the objective and constraint functionals. The optimization results will be affected by the selected initial geometry. However, the continuous solution to the problem which is developed here by means of two different approaches will be valid for any arbitrary initial geometry. Therefore, once it is successfully integrated with a numerical optimization method and an up-to-date flow solver such as SWAN, it can be arbitrarily applied in designs, and efficient and accurate hydrodynamic shape optimization results are obtained.

This thesis study is believed to provide a concise theoretical background for the shape optimization problem, which later needs to be implemented into numerical applications to obtain practical results. Another contribution aimed by this study is to relate the general optimal control theory with the Lagrange multipliers theorem for adjoint formulation so that the close relation of these complex concepts can be formed without being misguided by some one-sided and application-based literature available.

Although we constructed the problem in terms of Lagrange duality which is most likely to be convex independent of the original problem, uncertainties arising due to constraint qualifications need to be studied in detail and convex shape optimization formulations should be investigated in the future for a global solution to the problem.

Bibliography

[1] Boyd S., Vandenberghe L., *Convex Optimization,* Cambridge University Press, 2004.

[2] Ekeland I., Temam R., *Convex Anaylsis and Variational Problems,* North Holland Publishing Company Ltd., Amsterdam, 1976

[3] Ekeland I., Turnbull T., *Infinite Dimensional Optimization and Convexity,* The University of Chicago Press, Chicago, 1983.

[4] Giles, M.B., Pierce N.A., *An introduction to the Adjoint Approach to Design,* Flow, Turbulence and Combustion (65), 393-415, 2000.

[5] Kostyukov, A.A., *Theory of Ship Waves and Wave Resistance,* Effective Comm., Inc., Iowa City, Iowa, 1968.

[6] Hildebrand F.B., *Advanced Calculus for Applications,* Prentice-Hall, Inc., 1976.

[7] Luenberger, D.G., *Optimization by Vector Space Methods,* John Wiley & Sons, Inc., 1969.

[8] Lusternik, L.A., Sobolev, V.J., *Elements of Functional Analysis,* J.Wiley Publication, New York, 1974.

[9] Miller, A., *The isoperimetric Problem: A Motivated Look at the Calculus of Variations,* Mathematical Methods in Classical and Quantum Mechanics, 2000.

[10] Monge, F., Tobio, B., *Aerodynamic Design and Optimization by Means of Control Theory,* Computational Mechanics, 1998

[11] Newman J.N., *Marine Hydrodynamics,* MIT Press, Cambridge MA, 1977.

[12] Ragab, S.A., *An Adjoint Formulation for Shape Optimization in Free Surface Potential Flow,* Journal of Ship Research(45), 269-278, 2001.

[13] Rockafellar, R.T., *Conjugate Duality and Optimization,* Society of Industrial & Applied, 1974.

[14] Sagan, H., *Introduction to the Calculus of Variations,* New York, McGraw-Hill, 1969.

[15] Sclavounos, P., *Surface Waves and Their Interaction with Floating Bodies,* Lecture Notes, 13.022, MIT, 2003.

[16] Stoker J.J., *Water Waves: The Mathematical Theory with Applications,* Interscience Publishers, Inc., New York, 1957.

[17] Ta'asan, S., *Introduction to Shape Design and Control,* Von Karman Institute, 1997.